

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:08:34 ; Search time 1.12756 Seconds
(without alignments)
3666.126 Million cell updates/sec

Title: US-09-904-987-1

Perfect score: 222

Sequence: 1 DAEFRHDSGYEVHHQKLVFF.....VGSNKGAIIGLMVGVVIAT 43

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	222	100.0	57	E60045	Alzheimer's disease
2	222	100.0	57	F60045	Alzheimer's disease
3	222	100.0	57	D60045	Alzheimer's disease
4	222	100.0	57	D60045	Alzheimer's disease
5	222	100.0	57	A60045	Alzheimer's disease
6	222	100.0	57	B60045	Alzheimer's disease
7	222	100.0	82	P00438	Alzheimer's disease
8	222	100.0	695	A49795	Alzheimer's disease
9	222	100.0	770	Q10042	Alzheimer's disease
10	217	97.7	42	P00512	beta-amyloid prote
11	203	91.4	695	A27485	Alzheimer's disease
12	203	91.4	695	S00550	Alzheimer's disease
13	203	91.4	747	JH0773	Alzheimer's disease
14	133	58.9	33	S23094	beta-amyloid prote
15	64	28.8	755	A13228	tryptophan 2-monoo
16	62	27.9	755	1 DAAGWT	tryptophan 2-monoo
17	62	27.9	755	1 QAAG4T	genome polyprotein
18	57	25.7	327	S11435	general amino acid
19	57	25.7	503	S73843	SLC1 protein - yea
20	56.5	25.5	378	S61992	formylmethanofuran
21	55.5	25.0	297	D69525	Na+/H+ antiporter
22	55.5	25.0	621	AF3016	probable sodium/hy
23	55.5	25.0	642	B98268	3-methyl-2-oxobuta
24	55.5	25.0	678	G71526	glycosyl transfera
25	55	24.8	291	F95015	probable aldehyde
26	55	24.8	317	H97888	conserved hypothet
27	55	24.8	488	S27652	genome polyprotein
28	55	24.8	738	C95936	
29	55	24.8	3063	J50166	

ALIGNMENTS

RESULT 1

E60045

Alzheimer's disease amyloid beta/A4 protein precursor - sheep (fragment)

C:Species: Ovis sp. (sheep)

C:Date: 01-Dec-1992 #sequence_revision 01-Dec-1992 #text_change 28-Jul-1995

C:Accession: E60045

R:Johnstone, E.M.; Chaney, M.O.; Norris, F.H.; Pascual, R.; Little, S.P.

Brain Res. Mol. Brain Res. 10, 299-305, 1991

A:Title: Conservation of the sequence of the Alzheimer's disease amyloid peptide in d

A:Reference number: A60045; MUID:92017079; PMID:1656157

A:Accession: E60045

A:Molecule type: mRNA

A:Residues: 1-57 <JOH>

A:Cross-references: EMBL:X56130

C:Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type protein

C:Keywords: alternative splicing; Alzheimer's disease; amyloid; brain

Query Match 100.0%; Score 222; DB 2; Length 57;

Best Local Similarity 100.0%; Pred. No. 2.9e-22;

Matches 43; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGVVIAT 43

Db 6 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGVVIAT 48

RESULT 2

F60045

Alzheimer's disease amyloid beta/A4 protein precursor - pig (fragment)

C:Species: Sus scrofa domestica (domestic pig)

C:Date: 01-Dec-1992 #sequence_revision 01-Dec-1992 #text_change 13-Aug-1999

C:Accession: F60045

R:Johnstone, E.M.; Chaney, M.O.; Norris, F.H.; Pascual, R.; Little, S.P.

Brain Res. Mol. Brain Res. 10, 299-305, 1991

A:Title: Conservation of the sequence of the Alzheimer's disease amyloid peptide in d

A:Reference number: A60045; MUID:92017079; PMID:1656157

A:Accession: F60045

A:Molecule type: mRNA

A:Residues: 1-57 <JOH>

A:Cross-references: EMBL:X56127; NID:q1895; PIDN:CAA39592.1; PTD:g1896

C:Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type protein

C:Keywords: alternative splicing; Alzheimer's disease; amyloid; brain

Query Match 100.0%; Score 222; DB 2; Length 57;

Best Local Similarity 100.0%; Pred. No. 2.9e-22;

Matches 43; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGVVIAT 43

Db 6 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGVVIAT 48

A:Title: Homology of the amyloid beta protein precursor in monkey and human supports a P
A:Reference number: A49795; MUID:91273117; PMID:1905108
A:Accession: A49795
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-695 <POD>
A:Cross-references: GB:M58727; NID:g342062; PIDN:AAA36829.1; PID:g342063
C:Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinase i
C:Keywords: alternative splicing

Query Match 100.0%; Score 222; DB 1; Length 695;
Best Local Similarity 100.0%; Pred. No. 4.5e-21;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAEFRHDSGYEYHOKLVFFAEVDGSKGAIIGLMVGWGVAT 43
|||||
Db 597 DAEFRHDSGYEYHOKLVFFAEVDGSKGAIIGLMVGWGVAT 639
|||||

RESULT 9
ORHU44
Alzheimer's disease amyloid beta protein precursor [validated] - human
N:Alternate names: Alzheimer's disease amyloid A4 protein; coagulation factor X1a inhibitor
N:Contains: amyloid beta protein long, plaque form; amyloid beta protein short, vascular
protein precursor splice form APP(770)
C:Species: Homo sapiens (man)
C:Date: 30-Jun-1987 #sequence-revision 28-Jul-1995 #text-change 15-Sep-2000
A:Accession: S02260; S05194; A32277; A33260; A35486; I39452; I39453; I59562; A44
4668; A28583; A29302; A60805; JLO038; S06121; A60335; A59011; A38384; S29076; S38252; S3
R:Lemaire, H.G.; Salbaum, J.M.; Multhaup, G.; Kang, J.; Bayne, R.M.; Unterbeck, A.; Bey
Nucleic Acids Res. 17, 517-522, 1989
A:Title: The PreA4(695) precursor protein of Alzheimer's disease A4 amyloid is encoded b
A:Reference number: S02260; MUID:89128427; PMID:2763775
A:Accession: S02260
A:Molecule type: DNA
A:Residues: 1-288, 'V', 365-770 <LEMI>
A:Cross-references: EMBL:X13466
A:Note: alternative splice form APP(695)
R:Lemaire, H.G.
submitted to the EMBL Data Library, November 1988
A:Reference number: S05194
A:Accession: S05194
A:Molecule type: DNA
A:Residues: 1-14, 'W', 17-288, 'V', 365-770 <LEW2>
A:Cross-references: EMBL:X13466; NID:g35598; PIDN:CAA31830.1; PID:g871360
A:Note: alternative splice form APP(695)
R:La Fauci, G.; Lahiri, D.K.; Salton, S.R.J.; Robakis, N.K.
Biochem. Biophys. Res. Commun. 159, 297-304, 1989
A:Title: Characterization of the 5'-end region and the first two exons of the beta-prote
A:Reference number: A32277; MUID:89165870; PMID:2538123
A:Accession: A32277
A:Molecule type: DNA
A:Residues: 1-75 <LAF>
A:Cross-references: GB:M24546; GB:M24547; NID:g341202; PIDN:AA013654.1; PID:g516074
R:Johnstone, E.M.; Chaney, M.O.; Moore, R.E.; Ward, K.E.; Norris, F.H.; Little, S.P.
Biochem. Biophys. Res. Commun. 163, 1248-1255, 1989
A:Title: Alzheimer's disease amyloid peptide is encoded by two exons and shows similarit
A:Reference number: A33260; MUID:89392030; PMID:2675837
A:Accession: A33260
A:Molecule type: DNA
A:Residues: 656-737 <JOH>
A:Cross-references: GB:M29270; NID:g178863; PIDN:AAA51768.1; PID:g178865
R:Prelli, F.; Levy, E.; van Duinen, S.G.; Bots, G.T.A.M.; Luyendijk, W.; Frangione, B.
Biochem. Biophys. Res. Commun. 170, 301-307, 1990
A:Title: Expression of a normal and variant Alzheimer's beta-protein gene in amyloid of
A:Reference number: A35486; MUID:90321244; PMID:2196878
A:Accession: A35486
A:Molecule type: DNA
A:Residues: 672-710 <PRE1>
A:Note: 693-Gln was found in DNA isolated from HCHWA-D patients
R:Yoshikai, S.I.; Sakaki, H.; Doh-ura, K.; Furuya, H.; Sakaki, Y.
Gene 87, 257-263, 1990
A:Title: Genomic organization of the human amyloid beta-protein precursor gene.

A:Reference number: I39451; MUID:90236318; PMID:2110105
A:Accession: I39452
A>Status: nucleic acid sequence not shown; translation not shown; translated from GB/
A:Molecule type: DNA
A:Residues: 1-770 <YOS1>
A:Cross-references: GB:M33112; NID:g178613; PIDN:AAB59502.1; PID:g178616
A:Accession: I39451
A>Status: nucleic acid sequence not shown; translation not shown; translated from GB/
A:Molecule type: DNA
A:Residues: 1-530, 'OWLMPEVIPAFAWEAKVGR' <YOS2>
A:Cross-references: GB:M34875; NID:g178608; PIDN:AAB59501.1; PID:g178615
R:Yoshikai, S.I.; Sakaki, H.; Doh-ura, K.; Furuya, H.; Sakaki, Y.
Gene 102, 291-292, 1991
A:Reference number: A59020; MUID:91340168; PMID:1908403
A:Contents: annotation; erratum
A:Note: revised physical map for reference I39451
R:Levy, E.; Carman, M.D.; Fernandez-Madrid, I.J.; Power, M.D.; Lieberburg, I.; van Du
Science 248, 1124-1126, 1990
A:Title: Mutation of the Alzheimer's disease amyloid gene in hereditary cerebral hemo
A:Reference number: I39453; MUID:90260663; PMID:2111384
A:Accession: I39453
A>Status: translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 656-737 <LEV>
A:Cross-references: GB:M37896; NID:g178618; PIDN:AAA51727.1; PID:g178620
A:Note: a mutation with 693-Gln is presented
R:Murrell, J.; Farlow, M.; Ghetti, B.; Benson, M.D.
Science 254, 97-99, 1991
A:Title: A mutation in the amyloid precursor protein associated with hereditary Alzhe
A:Reference number: I59562; MUID:92022553; PMID:1925564
A:Accession: I59562
A>Status: translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 689-716, 'F', 718-737 <MUR>
A:Cross-references: GB:S57665; NID:g236720; PIDN:AAB19991.1; PID:g236721
R:Kamino, K.; Orr, H.T.; Payami, H.; Wijsman, E.M.; Alonso, M.E.; Pulst, S.M.; Anders
arakis, S.E.; Korenberg, J.R.; Sharma, V.; Kukull, W.; Larson, E.; Heston, L.L.; Mart
Am. J. Hum. Genet. 51, 998-1014, 1992
A:Title: Linkage and mutational analysis of familial Alzheimer disease kindreds for t
A:Reference number: A44017; MUID:93035397; PMID:1415269
A:Accession: A44017
A:Molecule type: DNA
A:Residues: 687-692, 'G', 694-718 <KAM1>
A:Cross-references: GB:S45135; NID:g257377; PIDN:AAB23645.1; PID:g257378
A:Experimental source: familial Alzheimer disease family SB
A:Note: sequence extracted from NCBI backbone (NCBIP:115374)
A:Accession: B44017
A:Molecule type: DNA
A:Residues: 687-718 <KAM2>
A:Cross-references: GB:S45136; NID:g257379; PIDN:AAB23646.1; PID:g257380
A:Experimental source: familial Alzheimer disease family LIT
A:Note: sequence extracted from NCBI backbone (NCBIP:115376)
R:Kang, J.; Lemaire, H.G.; Unterbeck, A.; Salbaum, J.M.; Masters, C.L.; Grzeschik, K.
Nature 325, 733-736, 1987
A:Title: The precursor of Alzheimer's disease amyloid A4 protein resembles a cell-sur
A:Reference number: A03134; MUID:87144572; PMID:2881207
A:Accession: A03134
A:Molecule type: mRNA
A:Residues: 1-288, 'V', 365-770 <KAN>
A:Cross-references: GB:Y00264; NID:g28525; PIDN:CAA68374.1; PID:g28526
A:Note: alternative splice form APP(695)
R:Robakis, N.K.; Ramakrishna, N.; Wolfe, G.; Wisniewski, H.M.
Proc. Natl. Acad. Sci. U.S.A. 84, 4190-4194, 1987
A:Title: Molecular cloning and characterization of a cDNA encoding the cerebrovascula
A:Reference number: A29030; MUID:87231971; PMID:3035574
A:Accession: A29030
A:Molecule type: mRNA
A:Residues: 284-288, 'V', 365-646, 'E', 648-770 <ROB>
A:Cross-references: GB:M16765; NID:g178539; PIDN:AAA51722.1; PID:g178540
A:Note: the authors translated the codon GAG for residue 647 as Asp
R:Goldgaber, D.; Lerman, M.I.; McBride, O.W.; Saffioti, U.; Gajdusek, D.C.
Science 235, 877-880, 1987

A:Residues: 507-770 <ZAI>
A:Cross-references: GB:M18734; NID:g178572; PIDN:AAA51726.1; PID:g178573
A:Note: the authors translated the codon GAA for residue 599 as Gly, ACC for residue 8 as Val, GTG for residue 609 as Asn, AAT for residue 610 as Gly, and GGT for residue 611 as Val. The cited Genbank accession number, J03594, is not in release 101.0
R: Masters, C.L.; Multhaup, G.; Simms, G.; Pottgiesser, J.; Martins, R.N.; Beyreuther, R.

Query Match 100.0%; Score 222; DB 1; Length 770;
Best Local Similarity 100.0%; Pred. No. 5e-21; 0; Indels 0; Gaps 0;
Matches 43; Conservative 0; Mismatches 0;

QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
|||||
Db 672 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 714
|||||

RESULT 10
PN0512
beta-amyloid protein - guinea pig (fragment)
C:Species: Cavia porcellus (guinea pig)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 17-Mar-1999
C:Accession: PN0512
R:Shimohigashi, Y.; Matsumoto, H.; Takano, Y.; Saito, R.; Iwata, T.; Kamiya, H.; Ohno
Biochem. Biophys. Res. Commun. 193, 624-630, 1993
A:Title: Receptor-mediated specific biological activity of a beta-amyloid protein fra
A:Reference number: PN0512; MUID:93290653; PMID:7685598
A:Accession: PN0512
A:Molecule type: protein
A:Residues: 1-42 <SHI>
C:Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinas
C:Keywords: alternative splicing; amyloid

Query Match 97.7%; Score 217; DB 2; Length 42;
Best Local Similarity 100.0%; Pred. No. 9.3e-22;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIA 42
|||||
Db 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIA 42
|||||

RESULT 11
A27485
Alzheimer's disease amyloid beta/A4 protein homolog precursor - mouse
N:Alternate names: proteinase nexin II
C:Species: Mus musculus (house mouse)
C:Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 13-Aug-1999
C:Accession: A27485; S19727; I49485
R:Yamada, T.; Sasaki, H.; Furuya, H.; Miyata, T.; Goto, I.; Sakaki, Y.
Biochem. Biophys. Res. Commun. 149, 665-671, 1987
A:Title: Complementary DNA for the mouse homolog of the human amyloid beta protein pr
A:Reference number: A27485; MUID:88106489; PMID:3322280
A:Accession: A27485
A:Molecule type: mRNA
A:Residues: 1-695 <YAM>
A:Cross-references: GB:M18373; NID:g191568; PIDN:AAA37139.1; PID:g309085
A:Experimental source: Brain
R:de Strooper, B.; van Leuven, F.; van den Berghe, H.
Biochim. Biophys. Acta 1129, 141-143, 1991
A:Title: The amyloid beta protein precursor or proteinase nexin II from mouse is clos
A:Reference number: S19727; MUID:92096458; PMID:1756177
A:Accession: S19727
A:Molecule type: mRNA
A:Residues: 1-210; G', 212-220, 'S', 222-396, 'A', 398-402, 'T', 404-448, 'A', 450-695 <STR>
A:Cross-references: EMBL:X59379
R:Izumi, R.; Yamada, T.; Yoshikai, S.; Sasaki, H.; Hattori, M.; Sakaki, Y.
Gene 112, 189-195, 1992
A:Title: Positive and negative regulatory elements for the expression of the Alzheimer
A:Reference number: I49485; MUID:92205998; PMID:1555768
A:Accession: I49485
A:Status: translated from GB/EMBL/DBDJB
A:Molecule type: DNA
A:Residues: 1-19 <RES>

Matches 16; Conservative 6; Mismatches 9; Indels 6; Gaps 1;

OY 7 DSGYEVHHOKLVFFAEDVGSNKGAIIGLWVGWVIAT 43

Db 223 DSG-----RIGFFPEDVPKPKVAIIIGAGISGLVVAS 253

Search completed: January 28, 2003, 14:19:54
Job time : 4.12756 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:03:18 ; Search time 0.657744 Seconds
(without alignments)
2711.515 Million cell updates/sec

Title: US-09-904-987-1

Perfect score: 222

Sequence: 1 DA6FRHDSGYEVHQRKLVFF.....VGSNGKALGLMVGGVVIAT 43

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	222	100.0	57	A4_PIG	Q29023 sus scrofa
2	222	100.0	57	A4_URSMA	Q29149 ursus marit
3	222	100.0	58	A4_CANFA	Q28280 canis famil
4	222	100.0	58	A4-RABIT	Q28748 oryctolagus
5	222	100.0	58	A4-SHEEP	Q28757 ovis aries
6	222	100.0	59	A4-BOVIN	Q28053 bos taurus
7	222	100.0	751	A4-SAISC	Q95341 salmirl sci
8	222	100.0	770	A4-HUMAN	P05067 homo sapien
9	203	91.4	770	A4-MOUSE	P12023 mus musculu
10	203	91.4	770	A4-RAT	P08592 rattus norv
11	62	27.9	755	TR2M_AGR4	P04029 agrobacteri
12	62	27.9	755	TR2N_AGRVI	P25017 agrobacteri
13	57	25.7	327	POLG_PVYCH	P21294 potatu viru
14	57	25.7	503	Y226_MYCPN	P75462 mycoplasma
15	56.5	25.5	378	SLGI_YEAST	P54867 saccharomyc
16	55.5	25.0	297	FTF_ARCFU	Q28076 archaeoglob
17	55	24.8	488	DHAL_PSESP	P33008 pseudomonas
18	55	24.8	3063	POLG_PVYN	P18247 p genome po
19	54.5	24.5	967	PDBL_ARATH	P93733 arabidopsis
20	54	24.3	284	POLG_PVYTO	P18977 potatu viru
21	54	24.3	322	Y853_RICPR	Q92ca7 rickettsia
22	53.5	24.1	708	YN2B_CAEEL	P45972 caenorhabdi
23	53.5	24.1	971	Y228_BORBU	O51246 borrelia bu
24	52	23.4	611	YCR3_YEAST	P25351 saccharomyc
25	51.5	23.2	1036	PLDI_CRIGR	O08684 cricetulus
26	51	23.0	494	COBQ_MYCTU	O53677 mycobacteri
27	51	23.0	763	APP2_HUMAN	Q06481 homo sapien
28	51	23.0	769	ITB2_BOVIN	P23592 bos taurus
29	51	23.0	3579	STAN_DROME	Q9v5n8 drosophila
30	50.5	22.7	915	PDB2_ARATH	Q23078 arabidopsis
31	50	22.5	285	1	P44860 haemophilus
32	50	22.5	674	1	P33543 arabidopsis
33	50	22.5	1437	1	O15440 homo sapien

34 49.5 22.3 1162 1 VGL2_IBVM
35 49 22.1 246 1 TPIS_CULTA
36 49 22.1 403 1 PGK_STRCO
37 49 22.1 432 1 PURA_RHILO
38 49 22.1 542 1 YM87_MYCTU
39 49 22.1 568 1 NIRS_PSEAE
40 49 22.1 695 1 APP2_MOUSE
41 49 22.1 701 1 YG5L_YEAST
42 49 22.1 704 1 SSP2_BOMMO
43 49 22.1 753 1 PPE2_HUMAN
44 49 22.1 757 1 PPE2_MOUSE
45 49 22.1 1080 1 HDC_DROME

ALIGNMENTS

RESULT 1

A4_PIG STANDARD; PRT; 57 AA.
AC Q29023;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Alzheimer's disease amyloid A4 protein homolog [Contains: Beta-amyloid
DE protein (Beta-APP) (A-beta)] (Fragment).
GN APP.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=92017079; PubMed=1656157;
RA "Johnstone E.M., Chaney M.O., Norris F.H., Pascual R., Little S.P.;
RT "Conservation of the sequence of the Alzheimer's disease amyloid
RT peptide in dog, polar bear and five other mammals by cross-species
RT polymerase chain reaction analysis."
RL Brain Res. Mol. Brain Res. 10:299-305(1991).
CC -!- FUNCTION: FUNCTIONAL NEURONAL RECEPTOR WHICH COUPLES TO
CC INTRACELLULAR SIGNALING PATHWAY THROUGH THE GTP-BINDING PROTEIN
CC G(O) (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: BELONGS TO THE APP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR HMBP; X56127; CAA39592.1; -.
DR HSSP; P05067; 1BA4.
DR InterPro; IPR001868; A4_APP.
DR InterPro; IPR001255; Beta-APP.
DR Fram; PF03494; Beta-APP; 1.
DR PROSITE; PS00319; A4_EXTRA; PARTIAL.
DR PROSITE; PS00320; A4_INTRA; PARTIAL.
KW Glycoprotein; Amyloid; Neutrone; Transmembrane.
FT NON_TER 1 1
FT CHAIN 6 48 BETA-AMYLOID PROTEIN (POTENTIAL).
FT DOMAIN <1 33 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 34 57 POTENTIAL.
FT NON_TER 57 57
SQ SEQUENCE 57 AA; 6172 MW; 84209D88EBA82DFA CRC64;

Query Match 100.0%; Score 222; DB 1; Length 57;
Best Local Similarity 100.0%; Pred. No. 2.7e-22;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Alzheimer's disease amyloid A4 protein homolog [Contains: Beta-amyloid protein (Beta-APP) (A-beta)] (Fragment).

APP.

Canis familiaris (Dog).

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.

NCBI_TaxID=9615;

[1]

SEQUENCE FROM N.A.

TISSUE=Kidney;

MEDLINE=92017079; PubMed=1656157;

Johnstone E.M., Chaney M.O., Norris F.H., Pascual R., Little S.P.; "Conservation of the sequence of the Alzheimer's disease amyloid peptide in dog, polar bear and five other mammals by cross-species polymerase chain reaction analysis.";

Brain Res. Mol. Brain Res. 10:299-305(1991).

!- FUNCTION: FUNCTIONAL NEURONAL RECEPTOR WHICH COUPLES TO INTRACELLULAR SIGNALING PATHWAY THROUGH THE GTP-BINDING PROTEIN G(O) (BY SIMILARITY).

!- SUBCELLULAR LOCATION: Type I membrane protein.

!- SIMILARITY: BELONGS TO THE APP FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X56125; CAA39590.1; -.

HSSP; P05067; IBA4.

InterPro; IPR001868; A4_APP.

InterPro; IPR001255; Beta-APP.

Pfam; PF03494; Beta-APP. 1.

PROSITE; PS00319; A4_EXTRA; PARTIAL.

PROSITE; PS00320; A4_INTRA; PARTIAL.

Glycoprotein; Amyloid; Neurone; Transmembrane.

NON_TER 1 1

CHAIN 7 49 BETA-AMYLOID PROTEIN (POTENTIAL).

DOMAIN <1 34 EXTRACELLULAR (POTENTIAL).

TRANSMEM 35 58 POTENTIAL.

NON_TER 58 58

SEQUENCE 58 AA; 6285 MW; 8469D488A2E12DFA CRC64;

Query Match 100.0%; Score 222; DB 1; Length 58;

Best Local Similarity 100.0%; Pred. No. 2.8e-22;

Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAEFRHDSGYEHHQKLVFFAEADVGSNGKAIGLMVGGVVIAT 43

DB 7 DAEFRHDSGYEHHQKLVFFAEADVGSNGKAIGLMVGGVVIAT 49

RESULT 4

A4_RABIT

ID A4_RABIT STANDARD; PRT; 58 AA.

AC Q28748;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Alzheimer's disease amyloid A4 protein homolog [Contains: Beta-amyloid protein (Beta-APP) (A-beta)] (Fragment).

GN APP.

OS Oryctolagus cuniculus (Rabbit).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.

OC NCBI_TaxID=9986;

OX [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Brain;

RC MEDLINE=92017079; PubMed=1656157;

RX Johnstone E.M., Chaney M.O., Norris F.H., Pascual R., Little S.P.;

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

EMBL; X56130; CAA39595.1; -
HSSP; P05067; IBA4.
InterPro; IPRO01868; A4_APP.
InterPro; IPRO01255; Beta_APP.
Pfam; PF03494; Beta_APP_1.
PROSITE; PS00319; A4_EXTRA; PARTIAL.
PROSITE; PS00320; A4_INTRA; PARTIAL.
Glycoprotein; Amyloid; Neurone; Transmembrane.
NON_TER 1 1
CHAIN 6 48 BETA-AMYLOID PROTEIN (POTENTIAL).
DOMAIN <1 33 EXTRACELLULAR (POTENTIAL).
TRANSMEM 34 57 POTENTIAL.
DOMAIN 58 >58 CYTOPLASMIC (POTENTIAL).
NON_TER 58 58
SEQUENCE 58 AA; 6300 MW; F434209D88EBA82D CRC64;

Query Match 100.0%; Score 222; DB 1; Length 58;
Best Local Similarity 100.0%; Pred. No. 2,8e-22;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNGKAIGLMVGGVVIAT 43
|||||
DB 6 DAEFRHDSGYEVHHQKLVFFAEDVGSNGKAIGLMVGGVVIAT 48

RESULT 6
A4_BOVIN

ID A4_BOVIN STANDARD; PRt; 59 AA.
AC Q28053;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Alzheimer's disease amyloid A4 protein homolog [Contains: Beta-amyloid
DE protein (Beta-APP) (A-beta)] (Fragment).
DE APP.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RS SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=92017079; PubMed=1656157;
RT Johnstone E.M., Chaney M.O., Norris F.H., Pascual R., Little S.P.;
RT "Conservation of the sequence of the Alzheimer's disease amyloid
RT peptide in dog, polar bear and five other mammals by cross-species
RT polymerase chain reaction analysis";
RL Brain Res. Mol. Brain Res. 10:299-305(1991).
CC !- FUNCTION: FUNCTIONAL NEURONAL RECEPTOR WHICH COUPLES TO
CC INTRACELLULAR SIGNALING PATHWAY THROUGH THE GTP-BINDING PROTEIN
CC G(O) (BY SIMILARITY).
CC !- SUBCELLULAR LOCATION: Type I membrane protein.
CC !- SIMILARITY: BELONGS TO THE APP FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

EMBL; X56124; CAA39589.1; -
EMBL; X56126; CAA39591.1; -

RA Untebeck A., Beyreuther K., Mueller-Hill B.;
 RT "The PreA4(695) precursor protein of Alzheimer's disease A4 amyloid
 RL is encoded by 16 exons.";
 RN Nucleic Acids Res. 17:517-522(1989).
 RP SEQUENCE FROM N.A.
 RX MEDLINE-97263807; PubMed-9108164;
 RA Hattori M., Tsukahara F., Furuhata Y., Tanahashi H., Hirose M.,
 RA Saito M., Tsukuni S., Sakaki Y.;
 RT "A novel method for making nested deletions and its application for
 RT sequencing of a 300 kb region of human APP locus.";
 RL Nucleic Acids Res. 25:1802-1808(1997).
 RN [5]
 RP SEQUENCE OF 286-345 AND 365-366 FROM N.A.
 RX MEDLINE-88122640; PubMed-2893290;
 RA Tanzi R.E., McClatchey A.I., Lamperti E.D., Villa-Komaroff L.,
 RA Gusella J.F., Neve R.L.;
 RT "Protease inhibitor domain encoded by an amyloid protein precursor
 RT mRNA associated with Alzheimer's disease.";
 RL Nature 331:528-530(1988).
 RN [6]
 RP SEQUENCE OF 287-367 FROM N.A.
 RX MEDLINE-88122641; PubMed-2893291;
 RA Kitaguchi N., Takahashi Y., Tokushima Y., Shiojiri S., Ito H.;
 RT "Novel precursor of Alzheimer's disease amyloid protein shows
 RT protease inhibitory activity.";
 RL Nature 331:530-532(1988).
 RN [7]
 RP SEQUENCE OF 284-289 AND 365-770 FROM N.A.
 RX MEDLINE-87231971; PubMed-3035574;
 RA Robakis N.K., Ramakrishna N., Wolfe G., Wisniewski H.M.;
 RT "Molecular cloning and characterization of a cDNA encoding the
 RT cerebrovascular and the neuritic plaque amyloid peptides.";
 RL Proc. Natl. Acad. Sci. U.S.A. 84:4190-4194(1987).
 RN [8]
 RP SEQUENCE OF 507-770 FROM N.A.
 RX MEDLINE-88124954; PubMed-2893379;
 RA Zain S.B., Salim M., Chou W.G., Sajdel-Sulkowska E.M., Majocha R.E.,
 RA Marotta C.A.;
 RT "Molecular cloning of amyloid cDNA derived from mRNA of the Alzheimer
 RT disease brain: coding and noncoding regions of the fetal precursor
 RT mRNA are expressed in the cortex.";
 RL Proc. Natl. Acad. Sci. U.S.A. 85:929-933(1988).
 RN [9]
 RP SEQUENCE OF 672-681.
 RX MEDLINE-88035004; PubMed-3312495;
 RA Partridge W.M., Vinters H.V., Yang J., Eisenberg J., Choi T.B.,
 RA Tourtellotte W.W., Huebner V., Shively J.E.;
 RT "Amyloid angiopathy of Alzheimer's disease: amino acid composition
 RT and partial sequence of a 4,200-dalton peptide isolated from cortical
 RT microvessels.";
 RL J. Neurochem. 49:1394-1401(1987).
 RN [10]
 RP SEQUENCE OF 739-770 FROM N.A.
 RX MEDLINE-90236318; PubMed-2110105;
 RA Yoshikawa S.-I., Sasaki H., Doh-Ura K., Furuya H., Sakaki Y.;
 RT "Genomic organization of the human amyloid beta-protein precursor
 RT gene.";
 RL Gene 87:257-263(1990).
 RN [11]
 RP SEQUENCE OF 1-10 FROM N.A.
 RX MEDLINE-89016647; PubMed-3140222;
 RA Schon E.A., Mita S., Sadlock J., Herbert J.;
 RT "A cDNA specifying the human amyloid beta precursor protein (ABPP)
 RT encodes a 95-kDa polypeptide.";
 RL Nucleic Acids Res. 16:9351-9351(1988).
 RN [12]
 RP SEQUENCE OF 18-50.
 RX MEDLINE-87250462; PubMed-3597385;
 RA van Nostrand W.E., Cunningham D.D.;
 RT "Purification of protease nexin II from human fibroblasts.";
 RL J. Biol. Chem. 262:8508-8514(1987).

RN [13]
 RP IDENTITY OF APP WITH NEXIN-II.
 RX MEDLINE-89384866; PubMed-2506449;
 RA Oltersdorf T., Fritz L.C., Schenk D.B., Lieberburg I.,
 RA Johnson-Wood K.L., Beattie E.C., Ward P.J., Blacher R.W., Dovey H.F.,
 RA Sinha S.;
 RT "The secreted form of the Alzheimer's amyloid precursor protein with
 RT the Kunitz domain is protease nexin-II.";
 RL Nature 341:144-147(1989).
 RN [14]
 RP PROTEASE-SPECIFICITY OF INHIBITOR DOMAIN.
 RX MEDLINE-90211252; PubMed-1969731;
 RA Kido H., Fukutomi A., Schilling J., Wang Y., Cordell B., Katunuma N.;
 RT "Protease-specificity of Kunitz inhibitor domain of Alzheimer's
 RT disease amyloid protein precursor.";
 RL Biochem. Biophys. Res. Commun. 167:716-721(1990).
 RN [15]
 RP COMPLEX WITH G(O).
 RX MEDLINE-93188965; PubMed-8446172;
 RA Nishimoto I., Okamoto T., Matsuura Y., Takahashi S., Okamoto T.,
 RA Murayama Y., Ogata E.;
 RT "Alzheimer amyloid protein precursor complexes with brain GTP-binding
 RT protein G(O).";
 RL Nature 362:75-79(1993).
 RN [16]
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 28-133.
 RX MEDLINE-99215582; PubMed-10201399;
 RA Rossjohn J., Cappell R., Feil S.C., Henry A., McKinstry W.J.,
 RA Galatis D., Hesse L., Multhaup G., Beyreuther K., Masters C.L.,
 RA Parker M.W.;
 RT "Crystal structure of the N-terminal, growth factor-like domain of
 RT Alzheimer amyloid precursor protein.";
 RL Nat. Struct. Biol. 6:327-331(1999).
 RN [17]
 RP X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS) OF 287-344.
 RX MEDLINE-91104913; PubMed-2125487;
 RA Hynes T.R., Randal M., Kennedy L.A., Eigenbrot C., Kossiakof A.A.;
 RT "X-ray crystal structure of the protease inhibitor domain of
 RT Alzheimer's amyloid beta-protein precursor.";
 RL Biochemistry 25:10018-10022(1990).
 RN [18]
 RP STRUCTURE BY NMR OF 289-344.
 RX MEDLINE-92031488; PubMed-1718421;
 RA Heald S.L., Tilton R.F. Jr., Hammond L.S., Lee A., Bayney R.M.,
 RA Kamack M.E., Ramabhadran T.V., Dreyer R.N., Davis G., Unterbeck A.,
 RA Tamburini P.P.;
 RT "Sequential NMR resonance assignment and structure determination of
 RT the Kunitz-type inhibitor domain of the Alzheimer's beta-amyloid
 RL precursor protein.";
 RL Biochemistry 30:10467-10478(1991).
 RN [19]
 RP STRUCTURE BY NMR OF 672-699.
 RX MEDLINE-94281210; PubMed-7516706;
 RA Talafous J., Marciniowski K.J., Klopman G., Zagorski M.G.;
 RT "Solution structure of residues 1-28 of the amyloid beta-peptide.";
 RL Biochemistry 33:7788-7796(1994).
 RN [20]
 RP STRUCTURE BY NMR OF 696-706.
 RX MEDLINE-97128622; PubMed-8973180;
 RA Kohno T., Kobayashi K., Maeda T., Sato K., Takashima A.;
 RT "Three-dimensional structures of the amyloid beta peptide (25-35) in
 RT membrane-mimicking environment.";
 RL Biochemistry 35:16094-16104(1996).
 RN [21]
 RP STRUCTURE BY NMR OF 672-711.
 RX MEDLINE-98359783; PubMed-9693002;
 RA Coles M., Bicknell W., Watson A.A., Fairlie D.P., Craik D.J.;
 RT "Solution structure of amyloid beta-peptide(1-40) in a water-miscible
 RT environment. Is the membrane-spanning domain where we think it is?";
 RL Biochemistry 37:11064-11077(1998).
 RN [22]
 RP STRUCTURE BY NMR OF 672-699.
 RX MEDLINE-20400066; PubMed-10940222;

RA Poulsen S.-A., Watson A.A., Craik D.J.;
 RT "Solution structures in aqueous SDS micelles of two amyloid beta
 RT peptides of Abeta(1-28) mutated at the alpha-secretase cleavage
 RT site";
 RL J. Struct. Biol. 130:142-152(2000).
 RN [23]
 RP STRUCTURE BY NMR OF 681-706.
 RX MEDLINE=20400065; PubMed=10940221;
 RA Zhang S., Iwata K., Lachenmann M.J., Peng J.W., Li S., Stimson E.R.,
 RA Lu Y., Felix A.M., Maggio J.E., Lee J.P.;
 RT "The Alzheimer's peptide a beta adopts a collapsed coil structure in
 RT water";
 RL J. Struct. Biol. 130:130-141(2000).
 RN [24]
 RP SIGNAL SEQUENCE CLEAVAGE SITE, AND TOPOLOGY.
 RX MEDLINE=88296437; PubMed=2500137;
 RA Dyrks J., Weidemann A., Multhaup G., Salbaum J.M., Lemaire H.-G.,
 RA Kang J., Mueller-Hill B., Masters C.L., Beyreuther K.;
 RT "Identification, transmembrane orientation and biogenesis of the
 RT amyloid A4 precursor of Alzheimer's disease";
 Query Match 100.08; Score 222; DB 1; Length 770;
 Best Local Similarity 100.08; Pred. No. 3.9e-21; Indels 0; Gaps 0;
 Matches 43; Conservative 0; Mismatches 0;
 QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
 DB 672 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 714
 RESULT 9
 A4_MOUSE
 ID A4_MOUSE STANDARD; PRT; 770 AA.
 AC PL2023;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Alzheimer's disease amyloid A4 protein homolog precursor
 DE (Amyloidogenic glycoprotein) (AG).
 GN APP.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 [1]
 RP SEQUENCE OF 1-289 AND 365-770 FROM N.A.
 RC STRAIN=BALB/C; TISSUE=Brain;
 RX MEDLINE=92096458; PubMed=1756177;
 RA de Strooper B., van Leuven F., van den Bergh H.;
 RT "The amyloid beta protein precursor or proteinase nexin II from mouse
 RT is closer related to its human homolog than previously reported";
 RL Biochim. Biophys. Acta 1129:141-143(1991).
 RN [2]
 RP SEQUENCE OF 1-289 AND 365-770 FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=88106489; PubMed=3322280;
 RA Yamada T., Sakaki H., Furuya H., Miyata T., Goto I., Sakaki Y.;
 RT "Complementary DNA for the mouse homolog of the human amyloid beta
 RT protein precursor";
 RL Biochem. Biophys. Res. Commun. 149:665-671(1987).
 RN [3]
 RP REVISIONS.
 RA Yamada T.;
 RL Submitted (MAR-1988) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 289-364 FROM N.A.
 RC STRAIN=CD-1; TISSUE=Placenta;
 RX MEDLINE=89345111; PubMed=2569710;
 RA Fukuchi K., Martin G.M., Deeb S.S.;
 RT "Sequence of the protease inhibitor domain of the A4 amyloid protein
 RT precursor of Mus domestica";
 RL Nucleic Acids Res. 17:5396-5396(1989).
 RN [5]

RP SEQUENCE OF 1-19 FROM N.A.
 RX MEDLINE=9220998; PubMed=1555768;
 RA Izumi R., Yamada T., Yoshikai S.I., Sakaki H., Hattori M.,
 RA Sakai Y.;
 RT "Positive and negative regulatory elements for the expression of the
 RT Alzheimer's disease amyloid precursor-encoding gene in mouse";
 RL Gene 112:189-195(1992).
 RN [6]
 RP SEQUENCE OF 281-380 FROM N.A., AND ALTERNATIVE SPLICING.
 RC TISSUE=Brain, and Kidney;
 RX MEDLINE=8919813; PubMed=2493250;
 RA Yamada T., Sakaki H., Dohura K., Goto I., Sakaki Y.;
 RT "Structure and expression of the alternatively-spliced forms of mRNA
 RT for the mouse homolog of Alzheimer's disease amyloid beta protein
 RT precursor";
 RL Biochem. Biophys. Res. Commun. 158:906-912(1989).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- ALTERNATIVE PRODUCTS: 5 ISOFORMS; APP(393), APP(563), APP(695),
 CC APP(751) AND APP(770) (SHOWN HERE); ARE PRODUCED BY ALTERNATIVE
 CC SPLICING.
 CC -1- TISSUE SPECIFICITY: AAA(770) IS EXPRESSED IN KIDNEY. AAA(751) IS
 CC WIDELY EXPRESSED. AAA(695) IS EXPRESSED IN BRAIN, KIDNEY AND
 CC LIVER.
 CC -1- DOMAIN: THE CLATHRIN-BINDING SITE IS ESSENTIAL FOR ITS ASSOCIATION
 CC WITH X11-ALPHA, -BETA, AND -GAMMA. THE SEQUENCE SPECIFIC
 CC RECOGNITION EXTENDS TO PEPTIDE RESIDUES THAT ARE C-TERMINAL TO THE
 CC NPXY MOTIF. THIS INTERACTION APPEARS TO BE INDEPENDENT OF
 CC PHOSPHORYLATION (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE APP FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 BPTI/KUNITZ INHIBITOR DOMAIN.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 DR EMBL; X59379; -; NOT_ANNOTATED_CDS.
 DR EMBL; M18373; AAA37139.1; -;
 DR EMBL; X15210; CAA33280.1; -;
 DR EMBL; D10603; BAA01456.1; -;
 DR EMBL; M24397; AAA33929.1; -;
 DR PIR; A27485; A27485.
 DR PIR; S04855; S04855.
 DR PIR; S19727; S19727.
 DR HSP; P05067; IAA.
 DR MGD; MGI:88059; App.
 DR InterPro; IPR001868; A4_APP.
 DR InterPro; IPR001255; Beta-APP.
 DR InterPro; IPR002223; Kunitz_BPTI.
 DR Pfam; PF00014; Kunitz_BPTI; 1.
 DR Pfam; PF02177; A4_EXTRA; 1.
 DR Pfam; PF03494; Beta-APP; 1.
 DR PRINTS; PR00203; AMYLOIDA4.
 DR PRINTS; PR00759; BASICPTASE.
 DR ProDom; PD000223; Kunitz_BPTI; 1.
 DR SMART; SM00006; A4_EXTRA; 1.
 DR SMART; SM00131; KU; 1.
 DR PROSITE; PS00319; A4_EXTRA; 1.
 DR PROSITE; PS00320; A4_INTRA; 1.
 DR PROSITE; PS00280; BPTI_KUNITZ_1; 1.
 DR PROSITE; PS00279; BPTI_KUNITZ_2; 1.
 KW Glycoprotein; Amyloid; Neuropeptide; Signal;
 KW Alternative splicing; Serine protease inhibitor.
 FT SIGNAL 1 17
 FT CHAIN 18 770
 FT ALZHEIMER'S DISEASE AMYLOID A4 PROTEIN
 FT HOMOLOG.
 FT DOMAIN 18 699
 FT EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 700 723
 FT POTENTIAL.
 FT DOMAIN 724 770
 FT CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 673 715
 FT EQUIVALENT OF BETA-AMYLOID PROTEIN.

```

FT DOMAIN 287 345 BPTI/KUNITZ INHIBITOR.
FT SITE 759 762 CLATHRIN-BINDING (BY SIMILARITY).
FT DISULFID 291 341 BY SIMILARITY.
FT DISULFID 300 324 BY SIMILARITY.
FT DISULFID 316 337 BY SIMILARITY.
FT CARBOHYD 542 542 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 571 571 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 289 364 E -> V (IN ISOFORM APP(695)).
FT VARSPLIC 290 364 MISSING (IN ISOFORM APP(695)).
FT VARSPLIC 346 380 MISSING (IN ISOFORM APP(751)).
SQ SEQUENCE 770 AA; 86752 MW; 26C50DE0890CAF7A CRC64;

Query Match 91.4%; Score 203; DB 1; Length 770;
Best Local Similarity 93.0%; Pred. No. 1.1e-18;
Matches 40; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 DAEFRHDSGYEVHOKLVFFAEADVGSNKGAIIGLMVGWGVVIAT 43
Db 672 DAEFGHDSGFVEVRHQKLVFFAEADVGSNKGAIIGLMVGWGVVIAT 714

RESULT 10
A4_RAT
ID A4_RAT STANDARD; PRT; 770 AA.
AC P08592;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Alzheimer's disease amyloid A4 protein homolog precursor
DE (Amyloidogenic glycoprotein) (AG).
GN APP.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE OF 1-289 AND 365-770 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=88312583; PubMed=2900758;
RA Shivers B.D., Hilbich C., Multhaup G., Salbaum J.M., Beyreuther K.,
RA Seeburg P.H.;
RT "Alzheimer's disease amyloidogenic glycoprotein: expression pattern
in rat brain suggests a role in cell contact.";
RL EMBO J. 7:1365-1370(1988).
RN [2]
RP SEQUENCE OF 289-364 FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=89183625; PubMed=2648331;
RA Kang J., Mueller-Hill B.;
RT "The sequence of the two extra exons in rat preA4.";
RL Nucleic Acids Res. 17:2130-2130(1989).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- ALTERNATIVE PRODUCTS: 5 ISOFORMS; APP(395), APP(563), APP(695),
CC APP(751) AND APP(770) (SHOWN HERE); ARE PRODUCED BY ALTERNATIVE
CC SPLICING.
CC -!- DOMAIN: THE CLATHRIN-BINDING SITE IS ESSENTIAL FOR ITS ASSOCIATION
CC WITH X11-ALPHA, -BETA, AND -GAMMA. THE SEQUENCE SPECIFIC
CC RECOGNITION EXTENDS TO PEPTIDE RESIDUES THAT ARE C-TERMINAL TO THE
CC NPXY MOTIF. THIS INTERACTION APPEARS TO BE INDEPENDENT OF
CC PHOSPHORYLATION (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE APP FAMILY.
CC -!- SIMILARITY: CONTAINS 1 BPTI/KUNITZ INHIBITOR DOMAIN.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; X07648; CAA30488.1; -.
EMBL; X14066; CAA32229.1; -.

```

```

DR PIR: S00550; S00550.
DR PIR: S03607; S03607.
DR HSP: P05067; IAAP.
DR InterPro: IPR001868; A4_APP.
DR InterPro: IPR001255; Beta-APP.
DR Pfam: PF00014; Kunitz_BPTI; 1.
DR Pfam: PF02177; A4_EXTRA; 1.
DR Pfam: PF03494; Beta-APP; 1.
DR PRINTS: PR00203; ANYLOIDA4.
DR PRINTS: PR00759; BASICPTASE.
DR PRODOM: PD000222; Kunitz_BPTI; 1.
DR SMART: SM00006; A4_EXTRA; 1.
DR SMART: SM00131; KU; 1.
DR PROSITE: PS00319; A4_EXTRA; 1.
DR PROSITE: PS00320; A4_INTRA; 1.
DR PROSITE: PS00280; BPTI_KUNITZ_1; 1.
DR PROSITE: PS00279; BPTI_KUNITZ_2; 1.
KW Glycoprotein; Amyloid; Neurone; Transmembrane; Signal;
KW Alternative splicing; Serine protease inhibitor.
FT SIGNAL 1 17 BY SIMILARITY.
FT CHAIN 18 770 ALZHEIMER'S DISEASE AMYLOID A4 PROTEIN
FT HOMOLOG.
FT DOMAIN 18 699 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 700 723 POTENTIAL.
FT DOMAIN 724 770 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 673 715 EQUIVALENT OF BETA-AMYLOID PROTEIN.
FT DOMAIN 287 345 BPTI/KUNITZ INHIBITOR.
FT SITE 759 762 CLATHRIN-BINDING (BY SIMILARITY).
FT DISULFID 291 341 BY SIMILARITY.
FT DISULFID 300 324 BY SIMILARITY.
FT DISULFID 316 337 BY SIMILARITY.
FT CARBOHYD 542 542 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 571 571 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 289 364 E -> V (IN ISOFORM APP(695)).
FT VARSPLIC 290 364 MISSING (IN ISOFORM APP(695)).
SQ SEQUENCE 770 AA; 86704 MW; C26C9D6BB2D929A7 CRC64;

Query Match 91.4%; Score 203; DB 1; Length 770;
Best Local Similarity 93.0%; Pred. No. 1.1e-18;
Matches 40; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 DAEFRHDSGYEVHOKLVFFAEADVGSNKGAIIGLMVGWGVVIAT 43
Db 672 DAEFGHDSGFVEVRHQKLVFFAEADVGSNKGAIIGLMVGWGVVIAT 714

RESULT 11
TR2M_AGR74
ID TR2M_AGR74 STANDARD; PRT; 755 AA.
AC P04029;
DT 23-OCT-1986 (Rel. 02, Created)
DT 23-OCT-1986 (Rel. 02, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tryptophan 2-monooxygenase (EC 1.13.12.3).
GN TWS1.
OS Agrobacterium tumefaciens (strain Ach5), and
OS Agrobacterium tumefaciens.
OG Plasmid pTiAch5, and plasmid pTiA6NC.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=176298, 358;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ach5; PLASMID=pTiAch5;
RX MEDLINE=84207942; PubMed=6327292;
RA Gielen J., de Beuckeleer M., Seurinck J., Deboeck F., de Greve H.,
RA Lemmers M., van Montagu M., Schell J.;
RT "The complete nucleotide sequence of the TL-DNA of the Agrobacterium
tumefaciens plasmid pTiAch5.";
RL EMBO J. 3:835-846(1984).
RN [2]
RP SEQUENCE FROM N.A.

```


Result No.	Score	Query			DB	ID	Description
		Match	Length				
1	222	100.0	82	4	Q16014	Q16014 homo sapien	
2	222	100.0	82	4	Q16019	Q16019 homo sapien	
3	222	100.0	82	4	Q16020	Q16020 homo sapien	
4	222	100.0	82	4	P78438	P78438 homo sapien	
5	222	100.0	534	13	O93296	O93296 gallus gall	
6	222	100.0	695	6	Q95KN7	Q95KN7 macaca fasc	
7	222	100.0	695	11	Q60496	Q60496 cavia sp. p	
8	222	100.0	695	13	Q9DGJ8	Q9DGJ8 gallus gall	
9	222	100.0	751	13	Q9DGJ7	Q9DGJ7 gallus gall	
10	222	100.0	770	6	Q9TU10	Q9TU10 sus scrofa	
11	217	97.7	569	13	Q9PVL1	Q9PVL1 gallus gall	
12	212	95.5	97	4	Q13778	Q13778 homo sapien	
13	203	91.4	79	11	O35463	O35463 cricetus	
14	203	91.4	607	11	Q99K32	Q99K32 mus musculus	
15	203	91.4	693	13	Q98SG0	Q98SG0 xenopus lae	
16	203	91.4	695	11	P97487	P97487 mus musculus	

```

AC Q16019;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Beta-amyloid peptide (Fragment).
GN BETA APP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=93236601; PubMed=8476439;
RX Denman R.B., Rosenzweig R., Miller D.L.;
RT "A system for studying the effect(s) of familial Alzheimer disease
RT mutations on the processing of the beta-amyloid peptide precursor.";
RL Biochem. Biophys. Res. Commun. 192:96-103(1993).
DR EMBL; S61380; AAB26264.2; -.
DR HSSP; P05067; 1BA4.
DR InterPro; IPR001255; Beta-APP.
DR Pfam; PF03494; Beta-APP; 1.
FT NON_TER 1
FT NON_TER 82
SQ SEQUENCE 82 AA; 8938 MW; F534AA50E579230A CRC64;

Query Match 100.0%; Score 222; DB 4; Length 82;
Best Local Similarity 100.0%; Pred. No. 1.3e-22;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
Db 18 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 60

RESULT 3
Q16020
ID Q16020;
AC Q16020;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Beta-amyloid peptide (Fragment).
GN BETA APP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=93236601; PubMed=8476439;
RX Denman R.B., Rosenzweig R., Miller D.L.;
RT "A system for studying the effect(s) of familial Alzheimer disease
RT mutations on the processing of the beta-amyloid peptide precursor.";
RL Biochem. Biophys. Res. Commun. 192:96-103(1993).
DR EMBL; S61383; AAB26265.2; -.
DR HSSP; P05067; 1BA4.
DR InterPro; IPR001255; Beta-APP.
DR Pfam; PF03494; Beta-APP; 1.
FT NON_TER 1
FT NON_TER 82
SQ SEQUENCE 82 AA; 8882 MW; F534AA5AE5D9230A CRC64;

Query Match 100.0%; Score 222; DB 4; Length 82;
Best Local Similarity 100.0%; Pred. No. 1.3e-22;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
Db 18 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 60

RESULT 4
P78438
ID P78438;
AC P78438;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Amyloid protein (Beta-amyloid protein) (Fragment).
GN APP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=89392030; PubMed=2675837;
RX Johnstone E.M., Chaney M.O., Moore R.E., Ward K.E., Norris F.H.,
RA Little S.P.;
RT "Alzheimer's disease amyloid peptide is encoded by two exons and shows
RT similarity to soybean trypsin inhibitor.";
RL Biochem. Biophys. Res. Commun. 163:1248-1255(1989).
RN [2]
RP SEQUENCE OF 19-48 FROM N.A.
RX MEDLINE=87120329; PubMed=2949367;
RA Tanzi R.E., Gusella J.F., Watkins P.C., Bruns G.A., George-Hyslop P.,
RA Van Keuren M.L., Patterson D., Pagan S., Kurnit D.M., Neve R.L.;
RT "Amyloid beta protein gene: cDNA, mRNA distribution, and genetic
RT linkage near the Alzheimer locus.";
RL Science 235:880-884(1987).
RN [3]
RP SEQUENCE OF 32-63 FROM N.A.
RX MEDLINE=93035397; PubMed=1415269;
RA Kanino K., Orr H.T., Payami H., Wijsman E.M., Alonso M.E., Pulst S.M.,
RA Anderson L., O'dahl S., Nemens E., White J.A.;
RT "Linkage and mutational analysis of familial Alzheimer disease
RT kindreds for the APP gene region.";
RL Am. J. Hum. Genet. 51:998-1014(1992).
DR EMBL; M29270; AAA51768.1; -.
DR EMBL; M29269; AAA51768.1; JOINED.
DR EMBL; M15532; AAA51564.1; -.
DR EMBL; S45136; AAB23646.1; -.
DR HSSP; P05067; 1BA4.
DR InterPro; IPR001255; Beta-APP.
DR Pfam; PF03494; Beta-APP; 1.
FT NON_TER 1
FT NON_TER 82
SQ SEQUENCE 82 AA; 8994 MW; 8DA9E42B813A070E CRC64;

Query Match 100.0%; Score 222; DB 4; Length 82;
Best Local Similarity 100.0%; Pred. No. 1.3e-22;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
Db 17 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 59

RESULT 5
Q93296
ID Q93296;
AC Q93296;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Amyloid protein (Fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98337885; PubMed=9671674;
RA Barnes N.Y., Li L., Yoshikawa K., Schwartz L.M., Oppenheim R.W.,
RA Milligan C.E.;
RT "Increased production of amyloid precursor protein provides a

```



```

RT substrate for caspase-3 in dying motoneurons.";
RL J. Neurosci. 18:5869-5880(1998).
DR EMBL; AF042098; AAC25052.1; -.
DR HSSP; P05067; 1BA4.
DR InterPro; IPR001868; A4_APP.
DR InterPro; IPR001255; Beta-APP.
DR Pfam; PF02177; A4_EXTRA; 1.
DR Pfam; PF03494; Beta-APP; 1.
DR PRINTS; PR00203; AMYLOIDA4.
DR PROSITE; PS00319; A4_EXTRA; 1.
DR PROSITE; PS00320; A4_INTRA; 1.
DR NON_TER 1
FT SEQUENCE 534 AA; 60597 MW; FB53ECC2E66D4C92 CRC64;
SQ
Query Match 100.0%; Score 222; DB 13; Length 534;
Best Local Similarity 100.0%; Pred. No. 1.2e-21;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
Db 436 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 478

RESULT 6
Q95KN7 PRELIMINARY; PRT; 695 AA.
AC Q95KN7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Amyloid b-protein precursor.
DE Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Macaca.
OC NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=CEREBELLUM;
RX MEDLINE=91273117; PubMed=1905108;
RA Podlisky M.B., Tolan D.R., Selkoe D.J.;
RT "Homology of the amyloid beta protein precursor in monkey and human
RT supports a primate model for beta amyloidosis in Alzheimer's
RT disease.";
RL Am. J. Pathol. 138:1423-1435(1991).
DR EMBL; M58727; AAA36829.1; -.
DR InterPro; IPR001868; A4_APP.
DR InterPro; IPR001255; Beta-APP.
DR Pfam; PF02177; A4_EXTRA; 1.
DR Pfam; PF03494; Beta-APP; 1.
DR PROSITE; PS00319; A4_EXTRA; 1.
DR PROSITE; PS00320; A4_INTRA; UNKNOWN_1.
FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 597 636 POTENTIAL.
SQ SEQUENCE 695 AA; 78663 MW; 4F6EA0139F969D56 CRC64;

Query Match 100.0%; Score 222; DB 6; Length 695;
Best Local Similarity 100.0%; Pred. No. 1.6e-21;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
Db 597 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 639

RESULT 7
Q60496 PRELIMINARY; PRT; 695 AA.
AC Q60496;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Putative amyloid precursor protein.

```

```

OS Cavia sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10143;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=97236426; PubMed=9116031;
RA Beck M., Mueller D., Bigl V.;
RT "Amyloid precursor protein in Guinea pigs - complete cDNA sequence and
RT alternative splicing.";
RL Biochim. Biophys. Acta 1351:17-21(1997).
DR EMBL; X97631; CAA66230.1; -.
DR HSSP; P05067; 1BA4.
DR InterPro; IPR001868; A4_APP.
DR InterPro; IPR001255; Beta-APP.
DR Pfam; PF02177; A4_EXTRA; 1.
DR Pfam; PF03494; Beta-APP; 1.
DR PRINTS; PR00203; AMYLOIDA4.
DR SMART; SM00006; A4_EXTRA; 1.
DR PROSITE; PS00319; A4_EXTRA; 1.
DR PROSITE; PS00320; A4_INTRA; 1.
DR PROSITE; PS00320; A4_INTRA; 1.
SQ SEQUENCE 695 AA; 78701 MW; 5196A0C4017F16AB CRC64;

Query Match 100.0%; Score 222; DB 11; Length 695;
Best Local Similarity 100.0%; Pred. No. 1.6e-21;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
Db 597 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 639

RESULT 8
Q9DGJ8 PRELIMINARY; PRT; 695 AA.
AC Q9DGJ8;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Beta-amyloid precursor protein 695 Isoform.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Sarasa M., Rodolosse A., Sorribas V.;
RT "Cloning of full-length chicken beta-amyloid precursor protein
RT isoforms.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF289218; AAG00593.1; -.
DR HSSP; P05067; 1BA4.
DR InterPro; IPR001868; A4_APP.
DR InterPro; IPR001255; Beta-APP.
DR Pfam; PF02177; A4_EXTRA; 1.
DR Pfam; PF03494; Beta-APP; 1.
DR PRINTS; PR00203; AMYLOIDA4.
DR SMART; SM00006; A4_EXTRA; 1.
DR PROSITE; PS00319; A4_EXTRA; 1.
DR PROSITE; PS00320; A4_INTRA; 1.
DR PROSITE; PS00320; A4_INTRA; 1.
SQ SEQUENCE 695 AA; 78565 MW; F201ED02AEC86D95 CRC64;

Query Match 100.0%; Score 222; DB 13; Length 695;
Best Local Similarity 100.0%; Pred. No. 1.6e-21;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
Db 597 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 639

```

```

RESULT 9
Q9DGJ7 PRELIMINARY; PRT; 751 AA.
ID Q9DGJ7
AC Q9DGJ7
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Beta-amyloid precursor protein 751 isoform.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Sarasa M., Rodolose A., Sorribas V.;
RT "Cloning of full-length chicken beta-amyloid precursor protein
RT isoforms.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF289219; AAC00594.1; -
DR HSSP; P05067; 1BA4.
DR InterPro; IPR001868; A4_APP.
DR InterPro; IPR001255; Beta-APP.
DR InterPro; IPR002223; Kunitz_BPTI.
DR Pfam; PF02177; A4_EXTRA; 1.
DR Pfam; PF03494; Beta-APP; 1.
DR Pfam; PF00014; Kunitz_BPTI; 1.
DR PRINTS; PR00203; AMYLOIDA4.
DR PRODOM; PD000222; Kunitz_BPTI; 1.
DR SMART; SM00131; KU; 1.
DR PROSITE; PS00319; A4_EXTRA; 1.
DR PROSITE; PS00320; BPTI_KUNITZ_1; 1.
DR PROSITE; PS0279; BPTI_KUNITZ_2; 1.
DR Serine protease inhibitor.
KW SEQUENCE 751 AA; 84705 MW; E78E9413A8033D84 CRC64;
SQ SEQUENCE 751 AA; 84705 MW; E78E9413A8033D84 CRC64;

Query Match 100.0%; Score 222; DB 13; Length 751;
Best Local Similarity 100.0%; Pred. No. 1.8e-21;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
DB 653 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 695

RESULT 10
Q9TU10 PRELIMINARY; PRT; 770 AA.
ID Q9TU10
AC Q9TU10
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Amyloid precursor protein.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Kimura A., Takahashi T.;
RT "Amyloid Precursor Protein 770.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB032550; BAA84580.1; -
DR HSSP; P05067; 1AAP.
DR InterPro; IPR001868; A4_APP.
DR InterPro; IPR001255; Beta-APP.
DR InterPro; IPR002223; Kunitz_BPTI.
DR Pfam; PF02177; A4_EXTRA; 1.
DR Pfam; PF03494; Beta-APP; 1.
DR Serine protease inhibitor.
KW SEQUENCE 770 AA; 86961 MW; 5F7A1DCB2BCC583E CRC64;
SQ SEQUENCE 770 AA; 86961 MW; 5F7A1DCB2BCC583E CRC64;

Query Match 97.7%; Score 217; DB 13; Length 569;
Best Local Similarity 100.0%; Pred. No. 6e-21;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIA 42
DB 472 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIA 513

RESULT 12
Q13778 PRELIMINARY; PRT; 97 AA.
ID Q13778
AC Q13778
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Amyloid protein (AD-AP) (Fragment).
OS Homo sapiens (Human).

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87120328; PubMed=3810169;
 RA Goldhaber D., Lerman M.I., McBride O.W., Saffioti U., Gajdusek D.C.;
 RT "Characterization and chromosomal localization of a cDNA encoding
 brain amyloid of Alzheimer's disease.";
 RL Science 235:877-880(1987).
 DR EMBL; M15533; AAA35540.1; -.
 DR HSSP; P05067; IBA4.
 DR InterPro; IPR001868; A4_APP.
 DR InterPro; IPR001255; Beta-APP.
 DR Pfam; PF03494; Beta-APP; 1.
 DR PRINTS; PR00203; AMYLOIDA4.
 FT NON_TER 1
 SQ SEQUENCE 97 AA; 10884 MW; E528CDB448DE474E CRC64;
 Query Match 95.5%; Score 212; DB 4; Length 97;
 Best Local Similarity 100.0%; Pred. No. 3.5e-21;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 EFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
 DB 1 EFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 41
 RESULT 13
 O35463
 ID O35463 PRELIMINARY; PRT; 79 AA.
 AC O35463;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Alzheimer's amyloid beta protein (Fragment).
 GN BETA APP.
 OS Cricetus griseus (Chinese hamster).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
 OC Cricetus
 OX NCBI_TaxID=10029;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Sambamurti K., Pinnix I., Gandhi S.;
 RL Submitted (Oct-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF030413; AAB86608.1; -.
 DR HSSP; P05067; IBA4.
 DR InterPro; IPR001255; Beta-APP.
 DR Pfam; PF03494; Beta-APP; 1.
 FT NON_TER 1
 FT NON_TER 79
 SQ SEQUENCE 79 AA; 8538 MW; 37F2C6C3BF3F597 CRC64;
 Query Match 91.4%; Score 203; DB 11; Length 79;
 Best Local Similarity 93.0%; Pred. No. 4.6e-20;
 Matches 40; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
 DB 21 DAEFGHDSGFVHRHOKLVFFAEDVGSNKGAIIGLMVGGVVAT 63
 RESULT 14
 Q99K32
 ID Q99K32 PRELIMINARY; PRT; 607 AA.
 AC Q99K32;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Hypothetical 68.4 kDa protein (Fragment).
 GN APP.
 OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC005490; AAH05490.1; -.
 DR HSSP; P05067; IAAAP.
 DR MGD; MGI:88059; App.
 DR InterPro; IPR001868; A4_APP.
 DR InterPro; IPR001255; Beta-APP.
 DR InterPro; IPR002223; Kunitz_BPTI.
 DR Pfam; PF02177; A4_EXTRA; 1.
 DR Pfam; PF03494; Beta-APP; 1.
 DR Pfam; PF00014; Kunitz_BPTI; 1.
 DR PRINTS; PR00203; AMYLOIDA4.
 DR PRINTS; PR00759; BASICPTASE.
 DR ProDom; PD000222; Kunitz_BPTI; 1.
 DR SMART; SM00131; KU; 1.
 DR PROSITE; PS00319; A4_EXTRA; 1.
 DR PROSITE; PS00320; A4_INTRA; 1.
 DR PROSITE; PS00280; BPTI_KUNITZ_1; 1.
 DR PROSITE; PS02079; BPTI_KUNITZ_2; 1.
 KW Hypothetical protein; Serine protease inhibitor.
 FT NON_TER 1
 SQ SEQUENCE 607 AA; 68391 MW; BF802214CBA7D172 CRC64;
 Query Match 91.4%; Score 203; DB 11; Length 607;
 Best Local Similarity 93.0%; Pred. No. 5.2e-19;
 Matches 40; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
 DB 509 DAEFGHDSGFVHRHOKLVFFAEDVGSNKGAIIGLMVGGVVAT 551
 RESULT 15
 Q98SG0
 ID Q98SG0 PRELIMINARY; PRT; 693 AA.
 AC Q98SG0;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Beta-amyloid precursor protein A.
 GN APP.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA van den Hurk W.H.;
 RL Thesis (2001), Department of Biological Sciences,
 University of Nijmegen, Nijmegen, Netherlands.
 DR EMBL; AJ298150; CAC37193.1; -.
 DR HSSP; P05067; IHZ3.
 DR InterPro; IPR001868; A4_APP.
 DR InterPro; IPR001255; Beta-APP.
 DR Pfam; PF02177; A4_EXTRA; 1.
 DR Pfam; PF03494; Beta-APP; 1.
 DR PRINTS; PR00203; AMYLOIDA4.
 DR SMART; SM00006; A4_EXTRA; 1.
 DR PROSITE; PS00319; A4_EXTRA; 1.
 DR PROSITE; PS00320; A4_INTRA; 1.
 KW SIGNAL.
 FT SIGNAL.
 SQ SEQUENCE 693 AA; 78568 MW; CAF1DF655C1AB653 CRC64;
 Query Match 91.4%; Score 203; DB 13; Length 693;
 Best Local Similarity 88.4%; Pred. No. 6.1e-19;
 Matches 38; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 13:58:08 : Search time 2.39607 Seconds
(without alignments)
2391.326 Million cell updates/sec

Title: US-09-904-987-1

Perfect score: 222

Sequence: 1 DAEPFRDSGVEVHHQKLVF.....VGSNKGALICLMVGGVVIAT 43

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	222	100.0	43	15	AA54759
2	222	100.0	43	15	AA60367
3	222	100.0	43	15	AA61328
4	222	100.0	43	16	AA64165
5	222	100.0	43	17	AA69573
6	222	100.0	43	18	AA93371
7	222	100.0	43	19	AA17758
8	222	100.0	43	19	AA71378
9	222	100.0	43	19	AA51316
10	222	100.0	43	19	AA40129

11	222	100.0	43	20	AA42955	Beta-amyloid precu
12	222	100.0	43	20	AAW89362	Beta-amyloid pepti
13	222	100.0	43	21	AA527020	Beta-amyloid pepti
14	222	100.0	43	21	AA515372	Human beta-amyloid
15	222	100.0	43	21	AA521216	Beta-amyloid pepti
16	222	100.0	43	21	AA58390	Beta-amyloid pepti
17	222	100.0	43	21	AA556102	Natural beta amylo
18	222	100.0	43	22	AA507901	Beta-amyloid pepti
19	222	100.0	43	22	AA578791	Human beta amyloid
20	222	100.0	43	22	AA512508	Beta-amyloid pepti
21	222	100.0	43	22	AA584428	Partial sequence o
22	222	100.0	43	22	AA598986	Beta-amyloid pepti
23	222	100.0	43	22	AA581193	Beta-amyloid pepti
24	222	100.0	43	22	AA591778	Amyloid beta-prote
25	222	100.0	43	22	AA591811	Amyloid beta-prote
26	222	100.0	43	22	AA547108	Biotinylated beta-
27	222	100.0	43	22	AA548344	Beta-amyloid pepti
28	222	100.0	43	23	AA598701	Human amyloid beta
29	222	100.0	43	23	AA50862	Beta-amyloid pepti
30	222	100.0	43	23	AA505149	Beta amyloid pepti
31	222	100.0	45	16	AA64169	Variant beta amylo
32	222	100.0	47	20	AA581475	Variant beta amylo
33	222	100.0	48	22	AA537523	Synthetic amyloid
34	222	100.0	52	16	AA64166	Amyloid precursor
35	222	100.0	52	20	AA581476	Variant beta amylo
36	222	100.0	53	15	AA55695	Synthetic amyloid
37	222	100.0	53	16	AA64168	Sequence of uniden
38	222	100.0	54	21	AA532126	Variant beta amylo
39	222	100.0	55	22	AA511482	Amyloid-beta precu
40	222	100.0	57	21	AA510910	Human APP peptide
41	222	100.0	59	17	AA505375	Human amyloid prec
42	222	100.0	59	19	AA570863	Amyloid precursor
43	222	100.0	59	22	AA584425	Beta-amyloid precu
44	222	100.0	60	21	AA569701	Partial sequence o
45	222	100.0	63	19	AA44747	Beta-amyloid precu

ALIGNMENTS

RESULT 1

AA54759

ID AA54759 standard; peptide; 43 AA.

XX AA54759;

AC AA54759;

DT 30-NOV-1994 (first entry)

XX Beta amyloid peptide.

DE Beta amyloid peptide.

DE Beta amyloid peptide.

DE Beta amyloid peptide.

DE Beta amyloid peptide.

DE Beta amyloid peptide.

DE Beta amyloid peptide.

DE Beta amyloid peptide.

DE Beta amyloid peptide.

DE Beta amyloid peptide.

DE Beta amyloid peptide.

DE Beta amyloid peptide.

DE Beta amyloid peptide.

DE Beta amyloid peptide.

DE Beta amyloid peptide.

DE Beta amyloid peptide.

DE Beta amyloid peptide.

DE Beta amyloid peptide.

DE Beta amyloid peptide.

DE Beta amyloid peptide.

DE Beta amyloid peptide.

DE Beta amyloid peptide.

DE Beta amyloid peptide.

DE Beta amyloid peptide.

DE Beta amyloid peptide.

DE Beta amyloid peptide.

DE Beta amyloid peptide.

DE Beta amyloid peptide.

DE Beta amyloid peptide.

Beta amyloid peptide; amyloid plaques; Alzheimer's disease; lesion; brain; senility; dementia; detection; diagnosis.

Homo sapiens.

WO9410569-A.

11-MAY-1994.

01-SEP-1993; 93WO-US08264.

26-OCT-1992; 92US-0965972.

(SCHE/) SCHENK D B.

(SCHL/) SCHLOSSMACHER M G.

(SELK/) SELKOE D.

Schenk DB, Schlossmacher MG, Selkoe DJ, Seubert PA;

Vigo-pelfrey;

WPI; 1994-167654/20.

Detecting soluble beta-amyloid peptide concns. e.g. for

PT diagnosing and assessing progression of Alzheimer's disease - by
 PT exposing cultured cells to test cpd. to determine effect of cpd.
 XX on produced soluble beta-amyloid peptide

XX Disclosure; Page 10; 55pp; English.

XX Beta amyloid peptide is the principal chemical constituent of
 CC amyloid plaques, lesions found on the brains of Alzheimer's disease
 CC patients. The ability to detect beta amyloid peptide in fluid
 CC samples provides a means of diagnosing Alzheimer's disease.

XX Sequence 43 AA;

Query Match 100.0%; Score 222; DB 15; Length 43;
 Best Local Similarity 100.0%; Pred. No. 1.9e-25;
 Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
 Db 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43

RESULT 2

AAR60367.
 ID AAR60367 standard; peptide; 43 AA.

XX AAR60367;

XX 15-MAR-1995 (first entry)

XX Beta-amyloid (1-43).

XX Amyloid precursor protein; APP; Alzheimer's disease; beta-amyloid;
 KW anti-beta-amyloid antibody; diagnosis.

XX Homo sapiens.

XX WO9417197-A.

XX 04-AUG-1994.

XX 24-JAN-1994; 94WO-JP00089.

XX 25-JAN-1993; 93JP-0010132.

XX 05-FEB-1993; 93JP-0019035.

XX 16-NOV-1993; 93JP-0286985.

XX 28-DEC-1993; 93JP-0334773.

XX (TAKE) TAKEDA CHEM IND LTD.

XX Kitada C, Odaka A, Suzuki N;

XX WPI; 1994-264110/32.

XX Antibodies recognising specific parts of beta-amyloid - can be
 PT used for diagnosis of diseases implicating beta-amyloid, such as
 PT Alzheimer's disease

XX Disclosure; Page 83; 116pp; Japanese.

XX Antibodies which recognise specific subfragments of the beta-amyloid
 CC protein are claimed. Specifically, the antibodies (which are pref.
 CC monoclonal) recognise residues 1-16 and/or 1-28 from the N-terminal
 CC portion of beta-amyloid or they recognise residues 25-35 or 35-43
 CC from the C-terminal portion. The antibodies are useful for assaying
 CC beta-amyloid and its derivatives for diagnosis of Alzheimer's
 CC disease.

XX Sequence 43 AA;

Query Match 100.0%; Score 222; DB 15; Length 43;
 Best Local Similarity 100.0%; Pred. No. 1.9e-25;
 Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
 Db 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43

RESULT 3

AAR61328
 ID AAR61328 standard; Protein; 43 AA.

XX AAR61328;

XX 21-APR-1995 (first entry)

XX Amyloid beta-protein, APP.

XX Amyloid beta-protein; APP; Tau-protein kinase I enzyme; TPK-I;

KW Alzheimer's disease.

XX Synthetic.

XX EP616032-A.

XX 21-SEP-1994.

XX 01-MAR-1994; 94EP-0103057.

XX 02-MAR-1993; 93JP-0041160.

XX 22-MAR-1993; 93JP-0085143.

XX 02-AUG-1993; 93JP-0191246.

XX (TAKA/) TAKASHIMA A.

PA (MITU) MITSUBISHI KASEI CORP.

XX Hoshino T, Imahori K, Saito K, Sato S, Shiratsuchi A;

PI Takashima A;

XX WPI; 1994-287181/36.

XX Newly isolated tau-protein kinase I enzyme - with specificity for
 PT tau-protein providing means for prevention and treatment of
 PT Alzheimer's disease

XX Example 1; Page 22; 30pp; English.

XX Amyloid beta-protein (APP) is the main component of senile plaques
 CC in Alzheimer's disease. APP was used to demonstrate the protective
 CC properties of anti-sense TPK-I oligonucleotides (AAQ67461 and AAQ67462)
 CC on hippocampus cells dosed with APP.

XX Sequence 43 AA;

Query Match 100.0%; Score 222; DB 15; Length 43;
 Best Local Similarity 100.0%; Pred. No. 1.9e-25;
 Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
 Db 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43

RESULT 4

AAR64165
 ID AAR64165 standard; peptide; 43 AA.

XX AAR64165;

XX 02-AUG-1995 (first entry)

XX Beta amyloid protein.

XX beta amyloid protein; mutant; variant; detection; amyloid deposition;
 KW diagnosis; amyloidosis associated disease; Alzheimer's disease;

KW Down's syndrome.
 XX Synthetic.
 OS WO9428412-A.
 PN 08-DEC-1994.
 PD 27-MAY-1994; 94WO-US05809.
 XX 28-MAY-1993; 93US-0069010.
 PF (MIRI-) MIRIAM HOSPITAL.
 PR Majocha RE, Marotta CA;
 XX WPI; 1995-023013/03.
 DR
 XX Amyloid binding composition comprising labelled amyloid protein
 PT and carrier - useful for in vivo imaging of amyloid deposits, for
 PT diagnosing Alzheimer's disease and Down's Syndrome.
 PT
 PS Claim 5; Page 42; 58pp; English.
 XX
 CC AAR64165 shows the amino acid sequence of the beta amyloid protein. The
 CC protein binds amyloid and is useful for in vivo imaging of amyloid
 CC deposits and hence diagnosis of an amyloidosis-associated disease, such
 CC as Alzheimer's disease or Down's syndrome. AAR64165-69 show specific
 CC variants generated from this generic sequence with addition amino acids.
 XX
 XX Sequence 43 AA;
 SQ
 Query Match 100.0%; Score 222; DB 16; Length 43;
 Best Local Similarity 100.0%; Pred. No. 1.9e-25;
 Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIAI 43
 DB 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIAI 43
 RESULT 5
 AAR95673
 ID AAR95673 standard; Peptide; 43 AA.
 XX
 AC AAR95673;
 XX
 DT 24-FEB-1997 (first entry)
 XX
 DE A-beta protein (43 amino acid version).
 XX
 XX A-beta protein; beta-amyloid precursor protein; APP; Alzheimer's disease;
 KW diagnosis; monitor; amyloid plaque; senile.
 KW
 XX Homo sapiens.
 OS
 XX WO9615452-A1.
 PN 23-MAY-1996.
 XX
 XX 13-NOV-1995; 95WO-US14659.
 PF
 XX 07-APR-1995; 95US-0419008.
 PR
 PR 14-NOV-1994; 94US-0339141.
 XX
 XX (ATHE-) ATHENA NEUROSCIENCES INC.
 PA
 XX Barbour R, Schenk DB, Seubert PA, Vigo-pelfrey C;
 PI WPI; 1996-260003/26.
 XX
 DR Diagnosis and monitoring of Alzheimer's disease - by detecting
 XX abnormally low concentration of A-beta peptide extending beyond
 PT

PT amino acid 41 in cerebrospinal fluid
 XX
 PS Disclosure; Page 16; 57pp; English.
 XX
 CC The "amyloid-beta peptide" or A-beta is an approx. 4.2 kD protein which,
 CC in the brains of Alzheimer's disease, Down's Syndrome, HCHWA-D and some
 CC normal aged subjects, forms the subunit of amyloid filaments comprising
 CC the senile (amyloid) plaques and the amyloid deposits in small cerebral
 CC and meningeal blood vessels. A-beta is an approx. 39-43 amino acid
 CC fragment of a large membrane-spanning glycoprotein, referred to as the
 CC beta-amyloid precursor protein (APP), encoded by a gene on the long arm
 CC of human chromosome 21. Detecting the amt. of A-beta is useful in
 CC diagnosis and monitoring of Alzheimer's disease, when taken together
 CC with other clinical symptoms. The present sequence is a 43 amino acid
 XX version of A-beta.
 XX
 SQ Sequence 43 AA;
 Query Match 100.0%; Score 222; DB 17; Length 43;
 Best Local Similarity 100.0%; Pred. No. 1.9e-25;
 Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIAI 43
 DB 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIAI 43
 RESULT 6
 AAW93371
 ID AAW93371 standard; Protein; 43 AA.
 XX
 AC AAW93371;
 XX
 DT 28-MAY-1999 (first entry)
 XX
 DE Human beta-amyloid polypeptide.
 XX
 XX Cathepsin Y; human; beta-amyloid peptide; BAP; secretion; inhibitor;
 KW protease; carboxypeptidase; aliphatic C-terminal amino acid; diagnostic;
 KW glycoprotein; cerebral deposition; pathogenesis; Alzheimer's disease;
 KW detection; amyloid plaque; angiodopathy; brain; Trisomy 21; amyloidosis;
 KW Down's syndrome; hereditary cerebral haemorrhage.
 XX
 OS Homo sapiens.
 XX
 XX WO9639194-A1.
 PN 12-DEC-1996.
 XX
 PD 26-APR-1996; 96WO-US06211.
 PF
 XX 06-JUN-1995; 95US-0469362.
 PR
 PR 06-JUN-1995; 95US-0467607.
 PR
 PR 02-MAY-1997; 97US-0850392.
 XX
 XX (ATHE-) ATHENA NEUROSCIENCES INC.
 PA (ANDE-) ANDERSON J.
 PA (CHRY/) CHRYSLER S.
 PA (MCCO/) MCCONLOGUE L.
 PA (SINH/) SINHA S.
 PA (TATS/) TATSUNO G.
 PA (TUNG/) TUNG JS.
 XX
 XX Anderson J, Chrysler S, McConlogue L, Semko CMF;
 PI Sinha S, Tatsuno G, Tung JS;
 PI WPI; 1997-042872/04.
 DR
 XX Acylamino and acyl-peptido-amino alcohol and aldehyde derivs. -
 PT inhibit beta-amyloid peptide prodn. in cells, use in Alzheimer's
 PT disease, also prepn. of cathepsin Y and nucleic acid encoding for
 PT it.
 XX

PS Disclosure; Page 11; 90pp; English.

XX This invention describes the inhibition of beta-amyloid peptide
 CC production in cells is effected by administration of an acylamino or
 CC acylpeptidoamino alcohol or aldehyde derivative of formula
 CC R1(X)m-Y-NR-CH(R2)-CH(R3)n-R4 where R and R10 = H or 1-6C alkyl;
 CC or R and R2 together, and/or R10 and R3 together complete a 4-10C ring
 CC structure; R1 = (a) 1-4C alkyl substd. by 1-5 substituents chosen from 6-10C
 CC aryl (opt. substd. by 1-3 of 1-6C alkyl, 1-6C alkoxy, 6-10C aryl, 6-10C
 CC aryloxy, OH, cyano, halo and amino), 3-8C cycloalkyl or Het, in which the
 CC substd. alkyl gp. is opt. further substd. by 1 or 2 of OH; (b) 2-4C
 CC alkenyl substd. by 1-4 substituents, as for 1-4C alkyl above; (c) 6-10C aryl
 CC (opt. substd. by 1-3 substituents, as for 1-4C alkyl above), (d) fluorenyl or (e)
 CC Het; Het = 3-14C heterocyclyl contg. 1-3 N, O or S heteroatoms; R2 and R3
 CC = D- or L- amino acid side chains of at least 2C atoms, excluding prolyl
 CC side chain; R4 = COCH=N2, CH2OH, C=NOH or COR5; R5 = H, 1-6C alkyl (opt.
 CC contg. 1 or 2 halogen atoms), 1-6C alkoxy, NR6R7 or N(Me)OMe; R6 and R7
 CC = H or 1-6C alkyl; X = O, NR9 or S; R9 = H, 1-6C alkyl or 6-10C aryl; Y =
 CC CO or CS; m = 0 or 1; and n = 0-2; provided that (i) when R1 =
 CC 1-naphthyl, R2 = CHMe2 (L-isomer), R3 = benzyl (L-isomer), Y = CO, m =
 CC 0 and n = 1, then R4 (sic) is not N(Me)OMe; (ii) when R1 = CHPh2, R2 =
 CC p-benzoyloxybenzyl (L-isomer), Y = CO, and m = n = 0, then R4 (sic) is
 CC not N(Me)OMe; and (iii) when R1 = PhCH=CPh, Y = CO, R2 = benzyl
 CC (L-isomer) and m = n = 0, then R4 (sic) is not N(Me)OMe. Cathepsin Y is
 CC a 31 kD carboxypeptidase with particular activity for aliphatic
 CC C-terminal amino acids. It is involved in the generation of beta-amyloid
 CC protein (BAP) from its precursor glycoprotein. Cerebral deposition of BAP
 CC plays a key role in pathogenesis of Alzheimer's disease (AD), and its
 CC detection can precede openly apparent symptoms of AD by years or more.
 CC Amyloid plaques and angiopathy are also present in the brain in those
 CC having Trisomy 21 (Down's syndrome) and hereditary cerebral haemorrhage
 CC with amyloidosis of Dutch type (HCHWA-D). The inhibitors are of use in
 CC inhibiting such plaque deposition and in treating AD. The cathepsin Y can
 CC be used, by virtue of its carboxypeptidase activity, in screening for BAP
 CC prodn. inhibitors, and, more widely, as a general C-terminal protease in
 CC a variety of chemical and biological systems. Conversely, assay of
 CC cathepsin Y in body fluid can be used diagnostically to evaluate the
 CC future risk of developing AD. The nucleic acid can be used to probe
 CC specifically for presence of cathepsin Y RNA or DNA in tissues or cloned
 CC libraries. It can also be used for expression of recombinant cathepsin Y,
 CC by transfection of a host cell.

XX
 SQ Sequence 43 AA;

Query Match 100.0%; Score 222; DB 18; Length 43;
 Best Local Similarity 100.0%; Pred. No. 1.9e-25;
 Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAEFRHDSGYEVHKKLVFFAEDVGSNKGAIIGLMVGGVVIAT 43
 |||||
 Db 1 DAEFRHDSGYEVHKKLVFFAEDVGSNKGAIIGLMVGGVVIAT 43
 |||||

RESULT 7
 AAAY17758
 ID AAY17758 standard; peptide: 43 AA.
 XX
 AC AAY17758;
 XX
 DT 10-AUG-1999 (first entry)
 XX
 DE Beta-amyloid peptide.
 XX
 KW Beta-amyloid peptide; beta-amyloid precursor protein; APP;
 KW membrane spanning glycoprotein; Alzheimer's disease.
 XX
 OS Homo sapiens.
 XX
 PN WO982493-A2.
 XX
 PD 28-MAY-1998.
 XX
 PF 20-NOV-1997; 97WO-US18704.

XX
 SQ Sequence 43 AA;

Query Match 100.0%; Score 222; DB 19; Length 43;
 Best Local Similarity 100.0%; Pred. No. 1.9e-25;
 Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAEFRHDSGYEVHKKLVFFAEDVGSNKGAIIGLMVGGVVIAT 43
 |||||
 Db 1 DAEFRHDSGYEVHKKLVFFAEDVGSNKGAIIGLMVGGVVIAT 43
 |||||

RESULT 8
 AAW71378
 ID AAW71378 standard; peptide: 43 AA.
 XX
 AC AAW71378;
 XX
 DT 02-DEC-1998 (first entry)
 XX
 DE Beta-amyloid precursor protein epitope.
 XX
 KW Beta amyloid precursor protein; eukaryotic cell line;
 KW exogenous gene construction; identification; quantification;
 KW inhibitor; beta-amyloid processing; treatment; Alzheimer's disease;
 KW brain trauma; Downs syndrome.
 XX
 OS Homo sapiens.
 XX
 PN WO9837215-A1.
 XX
 PD 27-AUG-1998.
 XX
 PF 03-FEB-1998; 98WO-US01899.
 XX
 PR 31-JUL-1997; 97US-0904296.
 PR 24-FEB-1997; 97US-0804971.
 PR 02-APR-1997; 97US-0825737.
 XX
 PA (HMRI) HOECHST MARION ROUSSEL INC.
 PA (SCIO-) SCIOS INC.
 XX
 PI Cordell B, Huggins J, Mischak RP, Pruss R, Rautmann G;
 PI Scardina JM;
 XX
 DR WPI; 1998-495368/42.
 XX

XX
 PR 22-NOV-1996; 96US-0755334.
 XX
 PA (ATHE-) ATHENA NEUROSCIENCES INC.
 PA (ELIL) LILLY & CO ELI.
 XX
 PI Audia JE, Folmer BK, John V, Latimer LH, Nissen JS;
 PI Porter WJ, Thorsett ED, Wu J;
 XX
 WPI; 1998-322359/28.
 DR
 XX New and known aryl or hetero-aryl amino acid derivatives - useful
 PT for inhibiting beta-amyloid peptide release and/or its synthesis and
 PT treating Alzheimer's disease
 XX
 PS Disclosure; Page 19; 131pp; English.
 XX
 CC The present invention describes a composition comprising an inert carrier
 CC and an N-(aryl or heteroaryl) amino acid derivative e.g.
 CC N-(N-(3,4-dichlorophenyl)alanyl) valine methyl ester. The composition
 CC can be used for inhibiting beta-amyloid peptide release and/or its
 CC synthesis in a cell, for preventing the onset of Alzheimer's disease
 CC and for treating Alzheimer's disease in order to inhibit further
 CC deterioration. The dosage is 0.1-500 mg/kg/day orally, rectally,
 CC transdermally, subcutaneously or intravenously. The present sequence
 CC represents the beta-amyloid peptide.
 XX
 SQ Sequence 43 AA;

PT New eukaryotic cell lines - containing a gene construct containing a
PT beta-amyloid precursor protein encoding sequence, used for
PT identifying inhibitors of beta-amyloid processing
XX
PS Disclosure; Fig 5; 82pp; English.
XX
CC The present sequence represents an epitope derived from a beta amyloid
CC precursor protein, and recognised by antibodies BA#1, 108.1, 1702.1,
CC and 1101.1. The specification describes an eukaryotic cell line
CC having exogenous gene construction comprising a cytomegalovirus (CMV)
CC promoter, a strong ribosome binding site (RBS), a beta-amyloid
CC precursor protein (BAPP) nucleic acid sequence, a selectable marker,
CC and a poly-adenylation signal. The products can be used for identifying
CC or quantifying beta-amyloid protein and for identifying inhibitors of
CC beta-amyloid processing. They can be used to develop products for
CC treating amyloidoses, e.g. Alzheimer's disease, brain trauma or Downs
CC syndrome.
XX
XX
SQ Sequence 43 AA;
Query Match 100.0%; Score 222; DB 19; Length 43;
Best Local Similarity 100.0%; Pred. No. 1.9e-25;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
Db 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
RESULT 9
AAW51316
XX ID AAW51316 standard; peptide; 43 AA.
XX AC AAW51316;
XX DT 14-AUG-1998 (first entry)
XX DE Natural beta-amyloid peptide fragment.
XX KW Natural beta-amyloid peptide; aggregation; D-amino acid;
XX KW Alzheimer's disease; beta-amyloidosis.
XX OS Homo sapiens.
XX PN WO98086868-Al.
XX PD 05-MAR-1998.
XX PF 27-AUG-1997; 97WO-US15166.
XX PR 21-JUL-1997; 97US-0897342.
XX PR 27-AUG-1996; 96US-0703675.
XX PA (PRAE-) PRAECIS PHARM INC.
XX PI Arico-muendel CC, Chin J, Flindeis MA, Geffer ML;
XX PI Hayward NJ, Kelley M, Komar-panicucci S, Lee J, Molineaux S;
XX PI Musso G, Phillips K, Signer ER, Wakefield J;
XX PR WPI; 1998-216936/19.
XX
XX PT Peptide compounds which are preferably based on beta-amyloid
XX PT peptide(s) - are useful in treatment of disorders related to
XX PT beta-amyloidosis, especially Alzheimer's disease
XX
XX PS Disclosure; Page 8; 92pp; English.
XX
CC The invention relates to peptides that modulate natural beta-amyloid
CC peptide aggregation. The modulators of the invention comprise a peptide
CC preferably based on a beta-amyloid peptide, that is comprised entirely
CC of D-amino acids. Preferably the peptide comprises 3-5 D-amino acid
CC residues and includes at least two D-amino acid residues independently
CC selected from the group consisting of D-leucine, D-phenylalanine and

CC D-valine. Preferred amino-terminal modifying groups include cyclic,
CC heterocyclic, polycyclic and branched alkyl groups. Preferred
CC carboxy-terminal modifying groups include an amide group, an alkyl amide
CC group, an aryl amide group or a hydroxy group. The peptides may be used
CC to treat disorders associated with beta-amyloidosis, especially
CC Alzheimer's disease. They may also be used in methods for detecting the
CC presence of beta-amyloid peptides in biological samples. The present
CC sequence represents the 43 amino acid long form of natural beta-amyloid
XX peptide.
XX
SQ Sequence 43 AA;
Query Match 100.0%; Score 222; DB 19; Length 43;
Best Local Similarity 100.0%; Pred. No. 1.9e-25;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
Db 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
RESULT 10
AAW40129
XX ID AAW40129 standard; peptide; 43 AA.
XX AC AAW40129;
XX DT 03-JUN-1998 (first entry)
XX DE Human amyloid-beta peptide.
XX KW Amyloid-beta peptide; A-beta; beta-AP; senile plaque; angiopathy;
XX KW brain; membrane-spanning glycoprotein; beta-amyloid precursor protein;
XX KW APP; Chromosome 21; human; Alzheimers disease; AD; amyloid filament;
XX KW treatment; disease; Down's syndrome; hereditary cerebral haemorrhage.
XX OS Homo sapiens.
XX PN WO9748983-Al.
XX PD 24-DEC-1997.
XX PF 18-JUN-1997; 97WO-US10601.
XX PR 18-JUN-1996; 96US-0665649.
XX PA (ATHE-) ATHENA NEUROSCIENCES INC.
XX PA (BGHM) BRIGHAM & WOMENS HOSPITAL.
XX PI Citron M, Schenk D, Selkoe DJ, Seubert PA;
XX PR WPI; 1998-063287/06.
XX
XX PT Identifying compounds that alter cellular production of amyloid-beta
XX PT 42 fragment - in vitro or in transgenic animal models, potentially
XX PT useful for treatment of Alzheimer's and other amyloid deposition
XX PT diseases
XX
XX PS Disclosure; Page 15; 86pp; English.
XX
CC This sequence represents a human amyloid-beta peptide (A-beta) which is a
CC fragment of the of a large membrane-spanning glycoprotein referred to as
CC the beta-amyloid precursor protein (APP) encoded by a gene on the long
CC arm of chromosome 21. This peptide is also known as the beta-AP peptide
CC and forms the subunit of the amyloid filaments comprising senile
CC (amyloid) plaques and the amyloid deposits in small cerebral and
CC meningeal blood vessels (amyloid angiopathy). The A-beta peptide can be
CC a 39-43 amino acid fragment. This invention provides methods of screening
CC compounds for their ability to alter the production of the A-beta
CC peptide, which is composed of >41 amino acids, alone, or in combination
CC with the A-beta peptide composed of 40 amino acids or less. Such agents
CC that reduce the production of the A-beta peptide are potentially useful
CC for treatment of Alzheimers Disease or other diseases involving amyloid

cc deposition such as Down's syndrome, hereditary cerebral haemorrhage with
cc amyloidosis of Dutch type and advanced aging of the brain.

XX	Sequence	43 AA:
SO		

Query Match	100.0%;	Score 222;	DB 19;	Length 43;
Best Local Similarity	100.0%;	Pred. No. 1.9e-25;		
Matches 43:	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QV	Matches	43; conservative	0; mismatches
1	DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIAT	43	0

QY I DAET RHDSGIEVAHHQRLFFFAEDVGSNAGRIIGTMVGGVIRI 43

Db 1 DAETFRHDSGIEVHHOKLYFFFAEDVGSNKGATIGLMVGGVVIAT 43

DO I WAFER KNDSOI EYVHHQVKNLVEFENECWVSCNANSCWIIISCEWVVCVVIINI 4

RESULT 11
AAY42955

AA142953
ID AAY42955 standard; protein; 43 AA.

XX AAY42955 standard, protein, 43 m.
AC AAY42955;

AC AM142955,
XX
DT 02-JAN-2000 (first entry)

DI 02-JAN-2000 (first entry)
XX
DE Beta-amyloid precursor protein.

DE beta-amyloid precursor protein.
XX
KW Beta-amyloid precursor protein; APP; release

XX OS Unidentified.
 XX PN WO200052048-A1.
 XX PD 08-SEP-2000.
 XX PF 03-MAR-2000; 2000WO-US05574.
 XX PR 04-MAR-1999; 99US-0122736.
 XX PA (PRAE-) PRAECIS PHARM INC.
 XX PI Findeis MA, Phillips K, Olson GL, Self C;
 XX DR WPI; 2000-594168/56.
 XX XX Novel compounds that are useful as modulators of beta-amyloid peptide
 PT aggregation in treating amyloidosis, comprises D-amino acids -
 XX
 XX PS Disclosure; Page 9; 87pp; English.
 XX CC The present sequence is beta-amyloid peptide.. The present invention
 CC relates to peptides (see AAB27023-B27046) that modulate beta-amyloid
 CC peptide aggregation, and hence inhibit the neurotoxicity of beta-amyloid
 CC peptide. The beta-amyloid peptide modulators would be useful for
 CC treating disorders associated with beta-amyloidosis for e.g. Alzheimer's
 CC disease. The present sequence is derived from amyloid precursor protein
 CC via proteolysis. The gene for amyloid precursor protein maps to
 CC chromosome 21.
 XX SQ Sequence 43 AA;
 Query Match 100.0%; Score 222; DB 21; Length 43;
 Best Local Similarity 100.0%; Pred. No. 1.9e-25;
 Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DAEFRHDSGYEVHHQKLVFAEDVGSNKGAIIGLMVGGVVIAI 43
 DB 1 DAEFRHDSGYEVHHQKLVFAEDVGSNKGAIIGLMVGGVVIAI 43
 RESULT 14
 AAB15372
 ID AAB15372 standard; peptide; 43 AA.
 XX AC AAB15372;
 XX DT 17-JAN-2001 (first entry)
 XX DE Human beta-amyloid precursor protein A-beta fragment.
 XX KW Human; beta-amyloid precursor protein; beta-APP; Alzheimer's disease;
 KW chromosome 21; epitope.
 XX OS Homo sapiens.
 XX PN WO200042166-A2.
 XX PD 20-JUL-2000.
 XX PF 13-JAN-2000; 2000WO-US00872.
 XX PR 13-JAN-1999; 99US-0115749.
 XX XX (DUPO) DUPONT PHARM CO.
 XX PI Seiffert DA, Mitchell TJ;
 XX DR WPI; 2000-476049/41.
 XX XX A CDNA construct that encodes beta-amyloid precursor protein for
 PT identifying compounds which inhibit A-beta peptide release and/or

PT synthesis comprises an epitope tag within the A-beta sequence -
 XX
 XX PS Disclosure; Page 13-14; 42pp; English.
 XX CC The present sequence is A-beta fragment of the human beta-amyloid
 CC precursor protein (beta-APP). It was used in the production of peptides
 CC of the invention. The invention concerns the production, detection and
 CC characterisation of epitope-tagged beta-APP proteins, and their use in
 CC identifying modulators of beta-APP which can be used to treat diseases
 CC associated with an altered metabolism of the protein, in particular
 CC Alzheimer's disease.
 XX SQ Sequence 43 AA;
 Query Match 100.0%; Score 222; DB 21; Length 43;
 Best Local Similarity 100.0%; Pred. No. 1.9e-25;
 Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DAEFRHDSGYEVHHQKLVFAEDVGSNKGAIIGLMVGGVVIAI 43
 DB 1 DAEFRHDSGYEVHHQKLVFAEDVGSNKGAIIGLMVGGVVIAI 43
 RESULT 15
 AAB21216
 ID AAB21216 standard; peptide; 43 AA.
 XX AC AAB21216;
 XX DT 11-JAN-2001 (first entry)
 XX DE Beta-amyloid peptide.
 XX KW Beta-amyloid peptide; membrane-spanning glycoprotein; analgesic;
 KW nicotinic agonist; beta-amyloid precursor protein; APP;
 KW Alzheimer's disease.
 XX OS Homo sapiens.
 XX PN US6117901-A.
 XX PD 12-SEP-2000.
 XX PF 21-NOV-1997; 97US-0976179.
 XX PR 22-NOV-1996; 96US-0098551.
 XX PA (ATHE-) ATHENA NEUROSCIENCES INC.
 XX PA (ELIL) LILLY & CO ELI.
 XX PI Thorsett ED, Nissen JS, Wu J, Latimer LH, John V, Fang LY;
 PI Audia JE, Mabry TE;
 XX DR WPI; 2000-637551/61.
 XX XX Eliciting analgesic effect in mammal, e.g. human, involves
 PT administering an aryl substituted olefinic amine compound -
 XX PS Disclosure; Column 11; 32pp; English.
 XX CC The present sequence is the beta-amyloid peptide, which is part of
 CC a large membrane-spanning glycoprotein, referred to as the beta-amyloid
 CC precursor protein (APP). Aryl substituted olefinic amine (metanilcotine)
 CC compounds which inhibit beta-amyloid peptide release and/or its
 CC synthesis may be useful for treating Alzheimer's disease, both
 CC prophylactically and therapeutically.
 XX SQ Sequence 43 AA;
 Query Match 100.0%; Score 222; DB 21; Length 43;
 Best Local Similarity 100.0%; Pred. No. 1.9e-25;
 Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAERHDSGYEVHHQKLVFFAEDVGSNKGAIGIMVGWGIAT 43
|||||

Db 1 DAERHDSGYEVHHQKLVFFAEDVGSNKGAIGIMVGWGIAT 43
|||||

Search completed: January 28, 2003, 14:13:52
Job time : 6.39607 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:19:14 ; Search time 0.56378 Seconds
(without alignments)
1539.035 Million cell updates/sec

Title: US-09-904-987-1

Perfect score: 222

Sequence: 1 DAEFRHDSGYEVHHQKLVFF.....VGSNKGAIIGLMVGGWIAT 43

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 122226 seqs, 20178551 residues

Total number of hits satisfying chosen parameters: 122226

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA:*

1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*

2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*

3: /cgn2_6/ptodata/2/pubpaa/US05_NEW_PUB.pep.*

4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*

5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*

6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*

7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*

8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*

9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*

10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*

11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*

12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*

13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*

14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	222	100.0	43	US-10-076-708-7	Sequence 7, Appli
2	222	100.0	43	US-10-051-496-1	Sequence 1, Appli
3	222	100.0	43	US-09-280-966-1	Sequence 1, Appli
4	222	100.0	43	US-09-904-987-1	Sequence 1, Appli
5	222	100.0	43	US-09-808-037-3	Sequence 3, Appli
6	222	100.0	43	US-09-866-712-3	Sequence 3, Appli
7	222	100.0	43	US-09-972-475-1	Sequence 3, Appli
8	222	100.0	43	US-09-982-800-1	Sequence 1, Appli
9	222	100.0	43	US-09-895-443-1	Sequence 1, Appli
10	222	100.0	43	US-09-996-357-1	Sequence 1, Appli
11	222	100.0	43	US-09-992-994-1	Sequence 1, Appli
12	222	100.0	43	US-09-984-834-1	Sequence 1, Appli
13	222	100.0	43	US-10-041-605-1	Sequence 1, Appli
14	222	100.0	53	US-09-797-543-5	Sequence 5, Appli
15	222	100.0	53	US-10-016-717-1	Sequence 1, Appli
16	222	100.0	70	US-09-155-076-14	Sequence 14, Appli
17	222	100.0	100	US-09-794-975-4	Sequence 4, Appli
18	222	100.0	103	US-09-972-475-2	Sequence 2, Appli
19	222	100.0	103	US-09-895-443-2	Sequence 2, Appli

20	222	100.0	117	10	US-09-794-975-6	Sequence 6, Appli
21	222	100.0	117	10	US-09-823-153-2	Sequence 2, Appli
22	222	100.0	355	10	US-09-794-975-13	Sequence 13, Appli
23	222	100.0	695	10	US-09-794-927-10	Sequence 10, Appli
24	222	100.0	695	10	US-09-794-927-12	Sequence 12, Appli
25	222	100.0	695	10	US-09-794-927-14	Sequence 14, Appli
26	222	100.0	695	10	US-09-795-847-10	Sequence 10, Appli
27	222	100.0	695	10	US-09-795-847-12	Sequence 12, Appli
28	222	100.0	695	10	US-09-795-847-14	Sequence 14, Appli
29	222	100.0	695	10	US-09-794-743-10	Sequence 10, Appli
30	222	100.0	695	10	US-09-794-743-12	Sequence 12, Appli
31	222	100.0	695	10	US-09-794-743-14	Sequence 14, Appli
32	222	100.0	695	10	US-09-794-748-10	Sequence 10, Appli
33	222	100.0	695	10	US-09-794-748-12	Sequence 12, Appli
34	222	100.0	695	10	US-09-794-748-14	Sequence 14, Appli
35	222	100.0	695	10	US-09-794-925-10	Sequence 10, Appli
36	222	100.0	695	10	US-09-794-925-12	Sequence 12, Appli
37	222	100.0	695	10	US-09-794-925-14	Sequence 14, Appli
38	222	100.0	695	10	US-09-681-442-10	Sequence 10, Appli
39	222	100.0	695	10	US-09-681-442-12	Sequence 12, Appli
40	222	100.0	695	10	US-09-681-442-14	Sequence 14, Appli
41	222	100.0	695	10	US-09-149-718-2	Sequence 2, Appli
42	222	100.0	697	10	US-09-794-927-16	Sequence 16, Appli
43	222	100.0	697	10	US-09-794-927-18	Sequence 18, Appli
44	222	100.0	697	10	US-09-794-927-20	Sequence 20, Appli
45	222	100.0	697	10	US-09-795-847-16	Sequence 16, Appli

ALIGNMENTS

RESULT 1

US-10-076-708-7

; Sequence 7, Application US/10076708

; Patent No. US20020164657A1

; GENERAL INFORMATION:

; APPLICANT: Shatma, Satish

; APPLICANT: Rank, Kenneth

; TITLE OF INVENTION: Assays for Accessing Aa-Tau Aggregation

; FILE REFERENCE: 6322

; CURRENT APPLICATION NUMBER: US/10/076.708

; CURRENT FILING DATE: 2002-02-15

; NUMBER OF SEQ ID NOS: 7

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 7

; LENGTH: 43

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-076-708-7

Query Match 100.0%; Score 222; DB 9; Length 43;

Best Local Similarity 100.0%; Pred. No. 1.6e-23;

Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGWIAT 43

Db 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGWIAT 43

RESULT 2

US-10-051-496-1

; Sequence 1, Application US/10051496

; Publication No. US20020182660A1

; GENERAL INFORMATION:

; APPLICANT: Kei-Lai L. Fong

; TITLE OF INVENTION: N- and C-Terminus Specific Immunoassays for

; Full Length Beta-Amyloid Peptide - Abeta(1-40), Abeta(1-42), Abeta(1-43)

; NUMBER OF SEQUENCES: 5

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Kei-Lai L. Fong

; STREET: 1004 West 8th Avenue

; CITY: King of Prussia

STATE: Pennsylvania
COUNTRY: USA
ZIP: 19406
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.50 inch, 1.44MB storage
COMPUTER: IBM PC Compatibles
OPERATING SYSTEM: Windows
SOFTWARE: MS No. US20020182660alepad
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/051.496
FILING DATE: 18-Jan-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/784.854A
FILING DATE: 16-Feb-2001
APPLICATION NUMBER: 60/183.407
FILING DATE: 18-February-2000
ATTORNEY/AGENT INFORMATION:
NAME: Koenig, C. Frederick III
REGISTRATION NUMBER: 29,662
REFERENCE/DOCKET NUMBER: PBI-PT001.1
TELEPHONE: (215) 568-6400
TELEFAX: (215) 568-6499
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 43 Amino Acid
TYPE: Amino Acid
STRANDEDNESS: <Unknown>
TOPOLOGY: Linear
MOLECULE TYPE: Protein
HYPOTHETICAL: <Unknown>
ANTI-SENSE: <Unknown>
ORIGINAL SOURCE:
ORGANISM: <Unknown>
INDIVIDUAL ISOLATE: <Unknown>
CELL TYPE: <Unknown>
IMMEDIATE SOURCE:
LIBRARY: <Unknown>
CLONE: <Unknown>
PUBLICATION INFORMATION:
AUTHORS:
TITLE:
JOURNAL:
VOLUME:
ISSUE:
PAGES:
DATE:
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 43
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-051-496-1

Query Match 100.0%; Score 222; DB 9; Length 43;
Best Local Similarity 100.0%; Pred. No. 1.6e-23;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAEFRHDSGYEVHHQKLVFAEDVGSNKGALIGLMVGGVVIAT 43
Db 1 DAEFRHDSGYEVHHQKLVFAEDVGSNKGALIGLMVGGVVIAT 43

RESULT 3

US-09-280-966-1
Sequence 1, Application US/09280966
Patent No. US20010020097A1
GENERAL INFORMATION:
APPLICANT: JAMES E. AUDIA
BEVERLY K. FOLMER
VARGHESE JOHN
JEFFREY S. NISSEN
WARREN J. PORTER
EUGENE D. THORSETT
JING WU

TITLE OF INVENTION: N-(ARYL/HETEROARYLACETYL) AMINO
ACID ESTERS, PHARMACEUTICAL COMPOSITIONS
COMPRISING SAME, AND METHODS FOR INHIBITING
-AMYLOID PEPTIDE RELEASE AND/OR ITS
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis, LLP
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/280,966
FILING DATE: 30-Mar-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/976,191
FILING DATE: 21 NOV 1997
APPLICATION NUMBER: 60/077,175
FILING DATE: 22 NOV 1996
ATTORNEY/AGENT INFORMATION:
NAME: Swiss, Gerald F.
REGISTRATION NUMBER: 30,113
REFERENCE/DOCKET NUMBER: 002010-335
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-622-2300
TELEFAX: 650-622-2499
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 43 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-280-966-1

Query Match 100.0%; Score 222; DB 10; Length 43;
Best Local Similarity 100.0%; Pred. No. 1.6e-23;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAEFRHDSGYEVHHQKLVFAEDVGSNKGALIGLMVGGVVIAT 43
Db 1 DAEFRHDSGYEVHHQKLVFAEDVGSNKGALIGLMVGGVVIAT 43

RESULT 4

US-09-904-987-1
Sequence 1, Application US/09904987
Patent No. US20020037908A1
GENERAL INFORMATION:

APPLICANT: No. US20020037908Alactyl, Inc.
TITLE OF INVENTION: Methods and Compositions for Controlling Pathological and Prep
TITLE OF INVENTION: Protein Assembly or Aggregation
FILE REFERENCE: 42108/26146
CURRENT APPLICATION NUMBER: US/09/904,987
CURRENT FILING DATE: 2001-07-12
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn version 3.0
SEQ ID NO: 1
LENGTH: 43
TYPE: PRT
ORGANISM: homo sapiens
PUBLICATION INFORMATION:
DATABASE ACCESSION NUMBER: NCBI ENTREZ / QRHUA4
DATABASE ENTRY DATE: 2000-09-15
RELEVANT RESIDUES: (672)..(714)
US-09-904-987-1

```

; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 09/216,958
;   FILING DATE: December 21, 1998
; ATTORNEY/AGENT INFORMATION:
;   NAME: Lee Cheng
;   REGISTRATION NUMBER: 40,949
; REFERENCE/DOCKET NUMBER: 2001-0488/LC/00177
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (202)721-8200
;   TELEFAX: (202)721-8250
;   TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 43 amino acids
;   TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-866-712-3

Query Match      100.0%; Score 222; DB 10; Length
Best Local Similarity 100.0%; Pred. No. 1.6e-23;
Matches 43; Conservative 0; Mismatches 0; Indels

QY 1 DAEFRHDSGVEVHHQKLVFFFAEDVGSNKGAIIGLMVGGVVIAT 43
    |||||
Db 1 DAEFRHDSGVEVHHQKLVFFFAEDVGSNKGAIIGLMVGGVVIAT 43

RESULT 7
US-09-972-475-1
; Sequence 1, Application US/09972475
; Patent No. US20020098173A1
; GENERAL INFORMATION:
; APPLICANT: Findeis, Mark A. et al.
; TITLE OF INVENTION: Modulators of Amyloid Aggregation
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/972,475
; FILING DATE: 04-Oct-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/617,267
; FILING DATE: <Unknown>
; APPLICATION NUMBER: USSN 08/475,579
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: USSN 08/548,998
; FILING DATE: 27-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: DeConti, Giulio A.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: PPI-002CP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 43 amino acids
;   TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide

```

```
;
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-972-475-1

Query Match      100.0%; Score 222; DB 10; Length 43;
Best Local Similarity 100.0%; Pred. No. 1.6e-23;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
    |||
DB 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43

RESULT 8
US-09-992-800-1
; Sequence 1, Application US/09992800
; Patent No. US20020102261A1
; GENERAL INFORMATION:
; APPLICANT: Raso, Victor
; TITLE OF INVENTION: IMMUNOLOGICAL CONTROL OF BETA-AMYLOID LEVELS IN VIVO
; FILE REFERENCE: BERI-2006
; CURRENT APPLICATION NUMBER: US/09/992,800
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: 09/594,366
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/139,408
; PRIOR FILING DATE: 1999-06-16
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 43
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-992-800-1

Query Match      100.0%; Score 222; DB 10; Length 43;
Best Local Similarity 100.0%; Pred. No. 1.6e-23;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
    |||
DB 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43

US-09-996-357-1
; Sequence 1, Application US/09996357
; Patent No. US20020133001A1
; GENERAL INFORMATION:
; APPLICANT: Gefter, Malcolm L
; APPLICANT: Isreal, David I
; APPLICANT: Joyal, John L
; APPLICANT: Gosselin, Michael
; TITLE OF INVENTION: THERAPEUTIC AGENTS AND METHODS OF USE THEREOF FOR
; FILE REFERENCE: PPI-105
; CURRENT APPLICATION NUMBER: US/09/996,357
; CURRENT FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: 60/253,302
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/250,198
; PRIOR FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: 60/257,186
; PRIOR FILING DATE: 2000-12-20
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 43
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-996-357-1

Query Match      100.0%; Score 222; DB 10; Length 43;
Best Local Similarity 100.0%; Pred. No. 1.6e-23;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
    |||
DB 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43

RESULT 10
US-09-996-357-1
; Sequence 1, Application US/09996357
; Patent No. US20020133001A1
; GENERAL INFORMATION:
; APPLICANT: Gefter, Malcolm L
; APPLICANT: Isreal, David I
; APPLICANT: Joyal, John L
; APPLICANT: Gosselin, Michael
; TITLE OF INVENTION: THERAPEUTIC AGENTS AND METHODS OF USE THEREOF FOR
; FILE REFERENCE: PPI-105
; CURRENT APPLICATION NUMBER: US/09/996,357
; CURRENT FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: 60/253,302
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/250,198
; PRIOR FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: 60/257,186
; PRIOR FILING DATE: 2000-12-20
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 43
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-996-357-1

Query Match      100.0%; Score 222; DB 10; Length 43;
Best Local Similarity 100.0%; Pred. No. 1.6e-23;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
    |||
DB 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43

RESULT 11
US-09-992-994-1
; Sequence 1, Application US/09992994
```


; Patent No. US20020136718A1
; GENERAL INFORMATION:
; APPLICANT: Raso, Victor
; TITLE OF INVENTION: IMMUNOLOGICAL CONTROL OF BETA-AMYLOID LEVELS IN VIVO
; FILE REFERENCE: BBRI-2005
; CURRENT APPLICATION NUMBER: US/09/992,994
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: 09/594,366
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/139,408
; PRIOR FILING DATE: 1999-06-16
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 43
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-992-994-1

Query Match 100.0%; Score 222; DB 10; Length 43;
Best Local Similarity 100.0%; Pred. No. 1.6e-23;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
|||||
Db 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43

RESULT 12
US-09-984-834-1
; Sequence 1, Application US/09984834
; Patent No. US20020137743A1
; GENERAL INFORMATION:
; APPLICANT: JING WU
; EUGENE D. THORSETT
; JEFFREY S. NISSEN
; THOMAS E. MABRY
; LEE H. LATIMER
; VARGHESE JOHN
; LAWRENCE Y. FANG
; JAMES E. AUDIA
; TITLE OF INVENTION: N-(ARYL/HERETOARYLACETYL)
; AMINO ACID ESTERS, PHARMACEUTICAL
; COMPOSITIONS COMPRISING SAME, AND
; METHODS FOR INHIBITING BETA-AMYLOID
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker &
; Mathis, LLP
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version
; #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/984,834
; FILING DATE: 31-Oct-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/976,179
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Swiss, Gerald F.
; REGISTRATION NUMBER: 30,113
; REFERENCE/DOCKET NUMBER: 002010-050
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-622-2300

; TELEFAX: 650-622-2499
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 43 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-984-834-1

Query Match 100.0%; Score 222; DB 10; Length 43;
Best Local Similarity 100.0%; Pred. No. 1.6e-23;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
|||||
Db 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43

RESULT 13
US-10-041-605-1
; Sequence 1, Application US/10041605
; Patent No. US20020115223A1
; GENERAL INFORMATION:
; APPLICANT: Tanzi, Rudolph E.
; APPLICANT: Bush, Ashley I.
; TITLE OF INVENTION: An In Vitro System For Determining Formation of A Amyloid
; FILE REFERENCE: 0609.4100001
; CURRENT APPLICATION NUMBER: US/10/041,605
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 08/294,819
; PRIOR FILING DATE: 1994-08-26
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 43
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-041-605-1

Query Match 100.0%; Score 222; DB 12; Length 43;
Best Local Similarity 100.0%; Pred. No. 1.6e-23;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
|||||
Db 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43

RESULT 14
US-09-797-543-5
; Sequence 5, Application US/09797543
; Patent No. US20020072050A1
; GENERAL INFORMATION:
; APPLICANT: Hook, Vivian Y. H.
; TITLE OF INVENTION: SECRETASES RELATED TO ALZHEIMER'S DEMENTIA
; FILE REFERENCE: P-As 4579
; CURRENT APPLICATION NUMBER: US/09/797,543
; CURRENT FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 09/173,887
; PRIOR FILING DATE: 1998-10-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 53
; TYPE: PRT
; ORGANISM: 'Axial Seamount' polynoid polychaete
US-09-797-543-5

Query Match 100.0%; Score 222; DB 10; Length 53;
Best Local Similarity 100.0%; Pred. No. 2e-23;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIAT 43
Db 4 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIAT 46

RESULT 15

US-10-016-717-1
; Sequence 1, Application US/10016717
; Patent No. US20020132281A1
; GENERAL INFORMATION:
; APPLICANT: Hook, Vivian Y.H.
; TITLE OF INVENTION: SECRETSSES RELATED TO ALZHEIMER'S DEMENTIA
; FILE REFERENCE: P-AS 5031
; CURRENT APPLICATION NUMBER: US/10/016.717
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: US 09/173,887
; PRIOR FILING DATE: 1998-10-16
; PRIOR APPLICATION NUMBER: US 09/294,987
; PRIOR FILING DATE: 1999-04-20
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 53
; TYPE: PRT
; ORGANISM: mammalian
US-10-016-717-1

Query Match 100.0%; Score 222; DB 12; Length 53;
Best Local Similarity 100.0%; Pred. No. 2e-23;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIAT 43
Db 4 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIAT 46

Search completed: January 28, 2003, 14:40:03
Job time : 1.56378 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:10:57 : Search time 0.892652 Seconds
(without alignments)
1417.333 Million cell updates/sec

Title: US-09-904-987-1

Perfect score: 222

Sequence: 1 DAEFRHDSGEVHHQKLVFF.....VGSNKGAIIGLMVGVVIAT 43

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A.COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B.COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A.COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B.COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS.COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	222	100.0	43	1	US-08-235-400-1
2	222	100.0	43	1	US-08-437-067-1
3	222	100.0	43	1	US-08-302-808-6
4	222	100.0	43	1	US-08-079-511-1
5	222	100.0	43	1	US-08-467-607-1
6	222	100.0	43	2	US-08-404-831-1
7	222	100.0	43	2	US-08-602-264A-3
8	222	100.0	43	2	US-08-469-362-1
9	222	100.0	43	2	US-08-612-785B-1
10	222	100.0	43	2	US-08-475-579A-1
11	222	100.0	43	2	US-08-850-392-1
12	222	100.0	43	2	US-08-986-348-6
13	222	100.0	43	2	US-08-975-977-1
14	222	100.0	43	2	US-08-817-423-1
15	222	100.0	43	2	US-08-920-162A-1
16	222	100.0	43	3	US-08-461-018A-3
17	222	100.0	43	3	US-08-976-191-1
18	222	100.0	43	3	US-08-976-179-1
19	222	100.0	43	4	US-09-216-958-3
20	222	100.0	43	4	US-09-356-931-1
21	222	100.0	43	4	US-08-733-202-1
22	222	100.0	43	4	US-08-703-675C-1
23	222	100.0	43	4	US-09-390-692-1
24	222	100.0	43	4	US-08-617-267C-1
25	222	100.0	43	4	US-09-303-655-1
26	222	100.0	43	4	US-08-294-819-1
27	222	100.0	47	2	US-08-609-090-10

28	222	100.0	52	2	US-08-609-090-11	Sequence 11, Appl
29	222	100.0	53	4	US-09-173-887-5	Sequence 5, Appl
30	222	100.0	59	1	US-08-484-969-3	Sequence 3, Appl
31	222	100.0	59	1	US-08-472-627-3	Sequence 3, Appl
32	222	100.0	59	1	US-08-388-463-3	Sequence 3, Appl
33	222	100.0	63	1	US-08-462-859A-4	Sequence 4, Appl
34	222	100.0	63	1	US-08-123-659A-4	Sequence 4, Appl
35	222	100.0	63	1	US-08-464-247A-4	Sequence 4, Appl
36	222	100.0	63	1	US-08-422-333-3	Sequence 3, Appl
37	222	100.0	99	3	US-08-339-708A-4	Sequence 4, Appl
38	222	100.0	99	3	US-08-339-708A-4	Sequence 4, Appl
39	222	100.0	99	3	US-08-339-708A-4	Sequence 4, Appl
40	222	100.0	100	6	5187153-10	Sequence 8, Appl
41	222	100.0	100	6	520013-10	Patent No. 5187153
42	222	100.0	100	6	5223482-10	Patent No. 5223482
43	222	100.0	103	2	US-08-404-831-2	Sequence 2, Appl
44	222	100.0	103	2	US-08-612-785B-2	Sequence 2, Appl
45	222	100.0	103	2	US-08-475-579A-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-235-400-1
; Sequence 1, Application US/08235400
; Patent No. 5552426
; GENERAL INFORMATION:
; APPLICANT: Lunn, William H.
; APPLICANT: Monn, James A.
; APPLICANT: Zimmerman, Dennis M.
; TITLE OF INVENTION: METHODS FOR TREATING A PHYSIOLOGICAL
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center/1104
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: United States of America
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/235,400
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Gaylo, Paul J.
; REGISTRATION NUMBER: 36,808
; REFERENCE/DOCKET NUMBER: X-9507
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317) 276-0756
; TELEFAX: (317) 276-3861
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 43 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-235-400-1

Query Match 100.0%; Score 222; DB 1; Length 43;

Best Local Similarity 100.0%; Pred. No. 3.9e-27; Indels 0; Gaps 0;

Matches 43; Conservative 0; Mismatches 0;

QY 1 DAEFRHDSGEVHHQKLVFFAEDVGSNKGAIIGLMVGVVIAT 43

Db 1 DAEFRHDSGEVHHQKLVFFAEDVGSNKGAIIGLMVGVVIAT 43

RESULT 2
US-08-437-067-1
; Sequence 1, Application US/08437067
; Patent No. 5593846
; GENERAL INFORMATION:
; APPLICANT: Schenk, Dale B.
; APPLICANT: Selkoe, Dennis J.
; APPLICANT: Schlossmacher, Michael G.
; APPLICANT: Seubert, Peter A.
; APPLICANT: Vigo-Pelfrey, Carmen
; TITLE OF INVENTION: Methods and Compositions for
; TITLE OF INVENTION: Detection
; TITLE OF INVENTION: of Soluble Beta-Amyloid Peptide
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend
; STREET: One Market Plaza, Steuart Tower, Suite 2000
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/437,067
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/965,972
; FILING DATE: 26-OCT-1992
; APPLICATION NUMBER: US 07/911,647
; FILING DATE: 10-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Heslin, James M.
; REGISTRATION NUMBER: 29,541
; REFERENCE/DOCKET NUMBER: 15270-6-1
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 43 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-437-067-1
Query Match 100.0%; Score 222; DB 1; Length 43;
Best Local Similarity 100.0%; Pred. No. 3.9e-27;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
DB 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
RESULT 3
US-08-302-808-6
; Sequence 6, Application US/08302808
; Patent No. 5750349
; GENERAL INFORMATION:
; APPLICANT: SUZUKI, No. 5750349uhiro
; APPLICANT: ODAKA, Asano
; APPLICANT: KITADA, Chieko
; TITLE OF INVENTION: ANTIBODIES TO B-AMYLOIDS OR THEIR
; TITLE OF INVENTION: DERIVATIVES AND USE THEREOF
; NUMBER OF SEQUENCES: 14

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN
; STREET: 130 WATER STREET
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02019
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/302,808
; FILING DATE: 15-SEP-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP94/00089
; FILING DATE: 24-JAN-1994
; APPLICATION NUMBER: 010132/1993
; FILING DATE: 25-JAN-1993
; APPLICATION NUMBER: 019035/1993
; FILING DATE: 05-FEB-1993
; APPLICATION NUMBER: 286985/1993
; FILING DATE: 16-NOV-1993
; APPLICATION NUMBER: 334773/1993
; FILING DATE: 28-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: DAVID, RESNICK S.
; REGISTRATION NUMBER: 34,235
; REFERENCE/DOCKET NUMBER: 44631
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; TELEX: 200291 STRE
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 43 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
US-08-302-808-6
Query Match 100.0%; Score 222; DB 1; Length 43;
Best Local Similarity 100.0%; Pred. No. 3.9e-27;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
DB 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
RESULT 4
US-08-079-511-1
; Sequence 1, Application US/08079511
; Patent No. 5766846
; GENERAL INFORMATION:
; APPLICANT: Schenk, Dale B.
; APPLICANT: Selkoe, Dennis J.
; APPLICANT: Schlossmacher, Michael G.
; APPLICANT: Seubert, Peter A.
; APPLICANT: Vigo-Pelfrey, Carmen
; TITLE OF INVENTION: Methods and Compositions for Detection
; TITLE OF INVENTION: of Soluble Beta-Amyloid Peptide
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend
; STREET: One Market Plaza, Steuart Tower, Suite 2000

;; CITY: San Francisco
;; STATE: California
;; COUNTRY: USA
;; ZIP: 94105
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; OPERATING SYSTEM: IBM PC compatible
;; SOFTWARE: PC-DOS/MS-DOS
;; CURRENT APPLICATION DATA: Patent In Release #1.0, Version #1.25
;; APPLICATION NUMBER: US/08/079,511
;; FILING DATE: 19930617
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/965,972
;; FILING DATE: 26-OCT-1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Heslin, James M.
;; REGISTRATION NUMBER: 29,541
;; REFERENCE/DOCKET NUMBER: 15270-6-1
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 415-326-2400
;; TELEFAX: 415-326-2422
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 43 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-08-079-511-1

Query Match 100.0%; Score 222; DB 1; Length 43;
Best Local Similarity 100.0%; Pred. No. 3.9e-27;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
|||||
Db 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43

RESULT 5
US-08-467-607-1
;; Sequence 1, Application US/08467607
;; Patent No. 5783434
;; GENERAL INFORMATION:
;; APPLICANT: TUNG, JAY S.
;; APPLICANT: SINHA, SUKANTO
;; APPLICANT: MCCONLOGUE, LISA
;; APPLICANT: TATSUNO, GWEN
;; APPLICANT: ANDERSON, JOHN
;; APPLICANT: CHRYSLER, SUSANNA
;; TITLE OF INVENTION: NOVEL CATHETESIN AND METHODS AND
;; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION THEREOF
;; NUMBER OF SEQUENCES: 20
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: ATHENA NEUROSCIENCES
;; STREET: 800 F. Gateway Blvd.
;; CITY: South San Francisco
;; STATE: CA
;; COUNTRY: USA
;; ZIP: 94080
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent In Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/467,607
;; FILING DATE: 06-JUN-1995
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: DUVALL, JEAN M.

;; REGISTRATION NUMBER: 32,731
;; REFERENCE/DOCKET NUMBER: 002010-007
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 877-0900
;; TELEFAX: (415) 877-8370
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 43 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
US-08-467-607-1

Query Match 100.0%; Score 222; DB 1; Length 43;
Best Local Similarity 100.0%; Pred. No. 3.9e-27;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
|||||
Db 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43

RESULT 6
US-08-404-831-1
;; Sequence 1, Application US/08404831
;; Patent No. 5817626
;; GENERAL INFORMATION:
;; APPLICANT: Mark A. Findeis, Howard Benjamin, Marc B. Garnick,
;; APPLICANT: Malcolm L. Geftter, Arvind Hundal, Laura Kasman,
;; APPLICANT: Gary Musso, Ethan R. Signer, and James Wakefield
;; TITLE OF INVENTION: Modulators of {SYMBOL 98 \f "Symbol"}-Amyloid Aggre
;; NUMBER OF SEQUENCES: 3
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: LAHIVE & COCKFIELD
;; STREET: 60 State Street, Suite 510
;; CITY: Boston
;; STATE: Massachusetts
;; COUNTRY: USA
;; ZIP: 02109-1875
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent In Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/404,831
;; FILING DATE: Herewith
;; CLASSIFICATION: 514
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Hanley, Elizabeth A. (EAH)
;; REGISTRATION NUMBER: 33,505
;; REFERENCE/DOCKET NUMBER: PPI-002
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (617)227-7400
;; TELEFAX: (617)227-5941
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 43 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; FRAGMENT TYPE: internal
US-08-404-831-1

Query Match 100.0%; Score 222; DB 2; Length 43;
Best Local Similarity 100.0%; Pred. No. 3.9e-27;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
|||||

Db 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43

RESULT 7
US-08-602-264A-3
; Sequence 3, Application US/08602264A
; Patent No. 5837853
; GENERAL INFORMATION:
; APPLICANT: AKIHiko TAKASHIMA et al.
; TITLE OF INVENTION: PREVENTIVE OR THERAPEUTIC AGENTS FOR
; TITLE OF INVENTION: ALZHEIMER'S DISEASE, A SCREENING METHOD OF ALZHEIMER'S DISEASE
; TITLE OF INVENTION: tau-PROTEIN KINASE I ORIGINATED FROM HUMAN BEING (AS AMENDED)
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WENDEROOTH, LIND & PONACK
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; COUNTRY: D.C.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch,
; MEDIUM TYPE: 144 mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/602.264A
; FILING DATE: February 20, 1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/204,091
; FILING DATE: March 2, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE:
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 43 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-602-264A-3

Query Match 100.0%; Score 222; DB 2; Length 43;
Best Local Similarity 100.0%; Pred. No. 3.9e-27;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
|||||
Db 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43

RESULT 8
US-08-469-362-1
; Sequence 1, Application US/08469362
; Patent No. 5849711
; GENERAL INFORMATION:
; APPLICANT: TUNG, JAY S.
; APPLICANT: SINHA, SUKANTO
; APPLICANT: MC CONLOGUE, LISA
; APPLICANT: SEMKO, CHRISTOPHER M.F.
; TITLE OF INVENTION: NOVEL CATHEPSIN AND METHODS AND
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION THEREOF
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ATHENA NEUROSCIENCES
; STREET: 900 F. Gateway Blvd.

Query Match 100.0%; Score 222; DB 2; Length 43;
Best Local Similarity 100.0%; Pred. No. 3.9e-27;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
|||||
Db 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43

RESULT 9
US-08-612-785B-1
; Sequence 1, Application US/08612785B
; Patent No. 5854204
; GENERAL INFORMATION:
; APPLICANT: FIndels, Mark A. et al.
; TITLE OF INVENTION: Ab Peptides that Modulate b-Amyloid
; TITLE OF INVENTION: Aggregation
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,785B
; FILING DATE: Herewith
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 08/404,831
; FILING DATE: 14-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 08/475,579
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 08/548,998
; FILING DATE: 27-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: DeConti, Giulio A.

Query Match 100.0%; Score 222; DB 2; Length 43;
Best Local Similarity 100.0%; Pred. No. 3.9e-27;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
|||||
Db 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43

REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: PPI-002CP3
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 43 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
US-08-612-785B-1

Query Match 100.0%; Score 222; DB 2; Length 43;
Best Local Similarity 100.0%; Pred. No. 3.9e-27;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
|||||
Db 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43

RESULT 10
US-08-475-579A-1
Sequence 1, Application US/08475579A
Patent No. 5854215
GENERAL INFORMATION:

APPLICANT: Mark A. Findels et al.
TITLE OF INVENTION: Modulators of {SYMBOL 98 \f "Symbol"}-Amyloid Peptide Aggrega
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,579A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/404,831
FILING DATE: 14-MAR-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Kara, Catherine J.
REGISTRATION NUMBER: P41,106
REFERENCE/DOCKET NUMBER: PPI-002CP
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 43 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
US-08-475-579A-1

Query Match 100.0%; Score 222; DB 2; Length 43;
Best Local Similarity 100.0%; Pred. No. 3.9e-27;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
|||||

Db 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43

RESULT 11

US-08-850-392-1
Sequence 1, Application US/08850392
Patent No. 5858982
GENERAL INFORMATION:

APPLICANT: TUNG, JAY S.
APPLICANT: SINHA, SURANTO
APPLICANT: MCCONLOGUE, LISA
APPLICANT: SEMKO, CHRISTOPHER M.F.
TITLE OF INVENTION: NOVEL CATHESPIN AND METHODS AND
TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION THEREOF
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: ATHENA NEUROSCIENCES
STREET: 800 F. Gateway Blvd.
CITY: South San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94080

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/850,392
FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/469,362
FILING DATE: 06-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: DUVALL, JEAN M.

REGISTRATION NUMBER: 32,731
REFERENCE/DOCKET NUMBER: 002010-005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 877-0900
TELEFAX: (415) 877-8370
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 43 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-850-392-1

Query Match 100.0%; Score 222; DB 2; Length 43;
Best Local Similarity 100.0%; Pred. No. 3.9e-27;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
|||||

Db 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43

RESULT 12

US-08-986-948-6
Sequence 6, Application US/08986948
Patent No. 5955317
GENERAL INFORMATION:

APPLICANT: SUZUKI, No. 5955317uhiro
APPLICANT: ODAKA, Asano
APPLICANT: KITADA, Chieko
TITLE OF INVENTION: ANTIBODIES TO B-AMYLOIDS OR THEIR
TITLE OF INVENTION: DERIVATIVES AND USE THEREOF
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN
STREET: 130 WATER STREET
CITY: BOSTON

STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PacenIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/817.423
FILING DATE: 4-AUG-1997

RESULT 13
us-08-975-977-1
; Sequence 1, Application US/08975977
; Patent No. 5965614
; GENERAL INFORMATION:
; APPLICANT: JAMES E. AUDIA
; APPLICANT: BEVERLY K. FOLMER
; APPLICANT: VARGHESE JOHN
; APPLICANT: LEE H. LATIMER
; APPLICANT: JEFFREY S. NISSEN
; APPLICANT: JON K. REEL
; APPLICANT: EUGENE D. THORSETT
; APPLICANT: CELIA A. WHITESITT
; TITLE OF INVENTION: N-(ARYL/HETEROARYL) AMINO
; ACID ESTERS, PHARMACEUTICAL
; TITLE OF INVENTION: COMPOSITIONS COMPRISING SAME, AND
; TITLE OF INVENTION: METHODS FOR INHIBITING BETA-AMYLROID
; TITLE OF INVENTION: PEPTIDE RELEASE AND/OR ITS

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US94/11895
FILING DATE: 19-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: KIM, JUDITH U.
REGISTRATION NUMBER: 40,679
REFERENCE/DOCKET NUMBER: 0609.4110000/REF/JUK
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 43 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-817-423-1

Query Match 100.0%; Score 222; DB 2; Length 43;
Best Local Similarity 100.0%; Pred. No. 3.9e-27;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
|||||
Db 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43

RESULT 15

US-08-920-162A-1
Sequence 1, Application US/08920162A
Patent No. 5985242
GENERAL INFORMATION:

APPLICANT: Findeis, M. et al.
TITLE OF INVENTION: Modulators of {SYMBOL 98 \f "Symbol"}-Amyloid Peptide
TITLE OF INVENTION: Aggregation Comprising D-Amino Acids
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE & COCKFIELD
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/920.162A
FILING DATE: Herewith
CLASSIFICATION: 514

PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/548,998
FILING DATE: 27-OCT-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/616,081
FILING DATE: 14-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/703,675
FILING DATE: 27-AUG-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER:

FILING DATE: 21-JUL-1997
ATTORNEY/AGENT INFORMATION:
NAME: KARA, Catherine J.
REGISTRATION NUMBER: 41,106
REFERENCE/DOCKET NUMBER: PPI-016CP4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 43 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
US-08-920-162A-1

Query Match 100.0%; Score 222; DB 2; Length 43;
Best Local Similarity 100.0%; Pred. No. 3.9e-27;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
|||||
Db 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43

Search completed: January 28, 2003, 14:21:17
Job time : 2.89265 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:08:34 ; Search time 20.1912 Seconds
(without alignments)
3666.126 Million cell updates/sec

Title: US-09-904-987-2
Perfect score: 4058
Sequence: 1 MLPGLALLLAANTARALEV.....KMQQNGYENPTKFFEQMN 770

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_73.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query %	Score	Match	Length	DB	ID	Description
1	4058	100.0	770	1	QRHUA4		Alzheimer's diseases
2	3590.5	88.5	695	1	A49795		Alzheimer's diseases
3	3493.5	86.1	695	2	S00550		Alzheimer's diseases
4	3468.5	85.5	695	2	A27485		Alzheimer's diseases
5	3403.5	83.9	747	2	JH0773		Alzheimer's diseases
6	2446.5	60.3	484	4	A32761		hypothetical Alzhe
7	1981.5	48.8	763	2	A49321		amyloid beta (A4)
8	1971.5	48.6	765	2	S42880		amyloid precursor-
9	1956.5	48.2	751	2	A49974		beta-amyloid precu
10	1150.5	28.4	653	2	A46362		amyloid precursor
11	1112.5	27.4	511	2	JC1404		CDEI-box DNA-bind
12	785	19.3	686	2	T15795		hypothetical prote
13	737.5	18.2	886	2	A32758		beta-amyloid-like
14	706	17.4	246	2	S38344		CDEI-binding prote
15	501	12.3	100	2	A32282		Alzheimer's diseases
16	411	10.1	82	2	PQ0438		Alzheimer's diseases
17	403	9.9	76	2	S06678		Alzheimer's diseases
18	397	9.8	76	2	S03607		Alzheimer's diseases
19	383	9.4	76	2	S04855		Alzheimer's diseases
20	296.5	7.3	191	2	A35981		sperm membrane pro
21	283	7.0	57	2	B60045		Alzheimer's diseases
22	283	7.0	57	2	F60045		Alzheimer's diseases
23	283	7.0	57	2	D60045		Alzheimer's diseases
24	283	7.0	57	2	D60045		Alzheimer's diseases
25	283	7.0	57	2	B60045		Alzheimer's diseases
26	283	7.0	57	2	B60045		Alzheimer's diseases
27	256.5	6.3	111	2	S41082		amyloid precursor
28	217	5.3	42	2	PN0512		beta-amyloid prote
29	194.5	4.8	2225	2	T26063		hypothetical prote

ALIGNMENTS

RESULT 1

QRHUA4

Alzheimer's disease amyloid beta protein precursor [validated] - human

N:Alternate names: Alzheimer's disease amyloid A4 protein; coagulation factor XIa inh
N:Contains: amyloid beta protein long, plaque form; amyloid beta protein short, vascu
Protein precursor splice form APP(770)

C:Species: Homo sapiens (man)

C>Date: 30-Jun-1987 #sequence_revision 28-Jul-1995 #text_change 15-Sep-2000

C:Accession: S02260; S05194; A32277; A3260; A35486; I39452; I39451; I39453; I59562;

4668; A28583; A29302; A60805; J10038; S06121; A60355; A59011; A38384; S29076; S38252;

R:Lemaire, H.G.; Salbaum, J.M.; Multhaup, G.; Kang, J.; Bayney, R.M.; Unterbeck, A.;

Nucleic Acids Res. 17, 517-522, 1989

A:Title: The PreA4(695) precursor protein of Alzheimer's disease A4 amyloid is encode

A:Reference number: S02260; MUID:89128427; PMID:2783775

A:Accession: S02260

A:Molecule type: DNA

A:Residues: 1-288, 'V', 365-770 <LEM1>

A:Cross-references: EMBL:X13466

A:Note: alternative splice form APP(695)

R:Lemaire, H.G.

submitted to the EMBL Data Library, November 1988

A:Reference number: S05194

A:Accession: S05194

A:Molecule type: DNA

A:Residues: 1-14, 'VW', 17-288, 'V', 365-770 <LEM2>

A:Cross-references: EMBL:X13466; NID:g35598; PIDN:CAA31830.1; PID:g871360

A:Note: alternative splice form APP(695)

R:La Fauci, G.; Lahiri, D.K.; Salton, S.R.J.; Robakis, N.K.

Biochem. Biophys. Res. Commun. 159, 297-304, 1989

A:Title: Characterization of the 5'-end region and the first two exons of the beta-pr

A:Reference number: A32277; MUID:89165870; PMID:2538123

A:Accession: A32277

A:Molecule type: DNA

A:Residues: 1-75 <LAF>

A:Cross-references: GB:M24546; GB:M24547; NID:g341202; PIDN:AA013654.1; PID:g516074

R:Johnstone, E.M.; Chaney, M.O.; Moore, R.E.; Ward, K.E.; Norris, F.H.; Little, S.P.

Biochem. Biophys. Res. Commun. 163, 1248-1255, 1989

A:Title: Alzheimer's disease amyloid peptide is encoded by two exons and shows simila

A:Reference number: A33260; MUID:89392030; PMID:2675837

A:Accession: A33260

A:Molecule type: DNA

A:Residues: 656-737 <JOH>

A:Cross-references: GB:M29270; NID:g178863; PIDN:AA051768.1; PID:g178865

R:Prelli, F.; Levy, E.; van Duinen, S.G.; Bots, G.T.A.M.; Luyendijk, W.; Frangione, B

Biochem. Biophys. Res. Commun. 170, 301-307, 1990

A:Title: Expression of a normal and variant Alzheimer's beta-protein gene in amyloid

A:Reference number: A35486; MUID:90321244; PMID:2196878

A:Accession: A35486

A:Molecule type: DNA

A:Residues: 672-710 <PRE1>

A:Note: 693-Gln was found in DNA isolated from HCHWA-D patients

R:Yoshikai, S.I.; Sasaki, H.; Doh-ura, K.; Furuya, H.; Sakaki, Y.

Gene 87, 257-263, 1990
A:Title: Genomic organization of the human amyloid beta-protein precursor gene.
A:Reference number: I39451; MUID:90236318; PMID:2110105
A:Accession: I39452
A>Status: nucleic acid sequence not shown; translation not shown; translated from GB/EMBL
A:Molecule type: DNA
A:Residues: 1-770 <YOS1>
A:Cross-references: GB:M33112; NID:g178613; PIDN:AAB59502.1; PID:g178616
A:Accession: I39451
A>Status: nucleic acid sequence not shown; translation not shown; translated from GB/EMBL
A:Molecule type: DNA
A:Residues: 1-530, 'QWLMPVPAFWEAKVGR' <YOS2>
A:Cross-references: GB:M34875; NID:g178608; PIDN:AAB59501.1; PID:g178615
R:Yoshikai, S.I.; Sasaki, H.; Doh-ura, K.; Furuya, H.; Sakaki, Y.
Gene 102, 291-292, 1991
A:Reference number: A59020; MUID:91340168; PMID:1908403
A:Contents: annotation; erratum
A>Note: revised physical map for reference I39451
R:Levy, E.; Carman, M.D.; Fernandez-Madrid, I.J.; Power, M.D.; Lieberburg, I.; van Duine
Science 248, 1124-1126, 1990
A:Title: Mutation of the Alzheimer's disease amyloid gene in hereditary cerebral hemorrhage
A:Reference number: I39453; MUID:90260663; PMID:2111584
A:Accession: I39453
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 656-737 <LEV>
A:Cross-references: GB:M37896; NID:g178618; PIDN:AAA51727.1; PID:g178620
A>Note: a mutation with 693-Gln is presented
R:Murrell, J.; Farlow, M.; Ghetti, B.; Benson, M.D.
Science 254, 97-99, 1991
A:Title: A mutation in the amyloid precursor protein associated with hereditary Alzheimer
A:Reference number: I59562; MUID:92022553; PMID:1925564
A:Accession: I59562
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 689-716, 'F', 718-737 <MUR>
A:Cross-references: GB:S57665; NID:g236720; PIDN:AAB19991.1; PID:g236721
R:Ramino, K.; Orr, H.T.; Payami, E.M.; Alonzo, M.E.; Pulst, S.M.; Anderson, R.;
arakis, S.E.; Korenberg, J.R.; Sharma, V.; Kukull, W.; Larson, E.; Heston, L.L.; Martin,
Am. J. Hum. Genet. 51, 998-1014, 1992
A:Title: Linkage and mutational analysis of familial Alzheimer disease kindreds for the
A:Reference number: A44017; MUID:93035397; PMID:1415269
A:Accession: A44017
A:Molecule type: DNA
A:Residues: 687-718 <KAN2>
A:Cross-references: GB:S45135; NID:g257377; PIDN:AAB23645.1; PID:g257378
A:Experimental source: familial Alzheimer disease family SB
A>Note: sequence extracted from NCBI backbone (NCBIP:115374)
A:Accession: B44017
A:Molecule type: DNA
A:Residues: 687-718 <KAN2>
A:Cross-references: GB:S45136; NID:g257379; PIDN:AAB23646.1; PID:g257380
A:Experimental source: familial Alzheimer disease family LIT
A>Note: sequence extracted from NCBI backbone (NCBIP:115376)
A>Note: this sequence has a silent mutation
R:Kang, J.; Lemaire, H.G.; Unterbeck, A.; Salbaum, J.M.; Masters, C.L.; Grzeschik, K.H.;
Nature 325, 733-736, 1987
A:Title: The precursor of Alzheimer's disease amyloid A4 protein resembles a cell-surface
A:Reference number: A03134; MUID:87144572; PMID:2881207
A:Accession: A03134
A:Molecule type: mRNA
A:Residues: 1-288, 'V', 365-770 <KAN>
A:Cross-references: GB:Y00264; NID:g28525; PIDN:CAA68374.1; PID:g28526
A>Note: alternative splice form APP(695)
R:Robakis, N.K.; Ramakrishna, N.; Wolfe, G.; Wisniewski, H.M.
Proc. Natl. Acad. Sci. U.S.A. 84, 4190-4194, 1987
A:Title: Molecular cloning and characterization of a cDNA encoding the cerebrovascular A
A:Reference number: A29030; MUID:87231971; PMID:3035574
A:Accession: A29030
A:Molecule type: mRNA
A:Residues: 284-288, 'V', 365-646, 'E', 648-770 <ROB>
A:Cross-references: GB:M16765; NID:g178539; PIDN:AAA51722.1; PID:g178540
A>Note: the authors translated the codon GAG for residue 647 as Asp

R:Goldgaber, D.; Lerman, M.I.; McBride, O.W.; Saffiotti, U.; Gajdusek, D.C.
Science 235, 877-880, 1987
A:Title: Characterization and chromosomal localization of a cDNA encoding brain amylo
A:Reference number: A47584; MUID:87120328; PMID:3810169
A:Accession: A47584
A:Molecule type: mRNA
A:Residues: 674-756, 'S', 758-770 <GOL>
A:Cross-references: GB:M15533; NID:g178706; PIDN:AAA35540.1; PID:g178707
A:Experimental source: brain
R:Tanzi, R.E.; Gusella, J.F.; Watkins, P.C.; Bruns, G.A.P.; St George-Hyslop, P.; Van
Science 235, 880-884, 1987
A:Title: Amyloid beta protein gene: cDNA, mRNA distribution, and genetic linkage near
A:Reference number: A47585; MUID:87120329; PMID:2949367
A:Accession: A47585
A:Molecule type: mRNA
A:Residues: 674-703 <TANI>
A:Cross-references: GB:M15532; NID:g177957; PIDN:AAA51564.1; PID:g177958
R:Dykes, T.; Weidemann, A.; Multhaup, G.; Salbaum, J.M.; Lemaire, H.G.; Kang, J.; Mue
EMBO J. 7, 949-957, 1988
A:Title: Identification, transmembrane orientation and biogenesis of the amyloid A4 p
A:Reference number: S02638; MUID:88296437; PMID:2900137
A:Accession: S02638
A:Molecule type: mRNA
A:Residues: 672-678 <DYR>
R:Tanzi, R.E.; McClatchey, A.I.; Lamperti, E.D.; Villa-Komaroff, L.; Gusella, J.F.; N
Nature 331, 528-530, 1988
A:Title: Protease inhibitor domain encoded by an amyloid protein precursor mRNA assoc
A:Reference number: S00707; MUID:88122640; PMID:2893290
A:Accession: S00707
A:Molecule type: mRNA
A:Residues: 286-344, 'I', 365-366 <TAN2>
A:Cross-references: EMBL:X06982; NID:g28817; PIDN:CAA30042.1; PID:g292612
A:Experimental source: promyelocytic leukemia cell line HL60
A>Note: alternative splice form APP(751)
R:Ponte, P.; Gonzalez-Dewhitt, P.; Schilling, J.; Miller, J.; Hsu, D.; Greenberg, B.;
Nature 331, 535-537, 1988
A:Title: A new A4 amyloid mRNA contains a domain homologous to serine proteinase inh
A:Reference number: S00925; MUID:88122639; PMID:2893289
A:Accession: S00925
A:Molecule type: mRNA
A:Residues: 1-344, 'I', 365-770 <PO>
A:Cross-references: GB:X06989; EMBL:Y00297; NID:g28720; PIDN:CAA30050.1; PID:g28721
A>Note: alternative splice form APP(751)
R:Kitaguchi, N.; Takahashi, Y.; Tokushima, Y.; Shiojiri, S.; Ito, H.
Nature 331, 530-532, 1988
A:Title: Novel precursor of Alzheimer's disease amyloid protein shows protease inhibi
A:Reference number: A38949; MUID:88122641; PMID:2893291
A:Accession: A38949
A:Molecule type: mRNA
A:Residues: 287-367 <KIT>
A:Cross-references: GB:X06981; NID:g28816; PIDN:CAA30041.1; PID:g292611
A:Experimental source: glioblastoma cell line
A>Note: alternative splice form APP(770)
R:Vitek, M.P.; Rasool, C.G.; de Sauvage, F.; Vitek, S.M.; Bartus, R.T.; Beer, B.; Ash
Brain Res. Mol. Brain Res. 4, 121-131, 1988
A:Title: Absence of mutation in the beta-amyloid cDNAs cloned from the brains of three
A:Reference number: A30320
A:Accession: A30320
A>Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 284-288, 'V', 365-770 <VIT1>
A:Accession: B30320
A>Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 122-288, 'V', 365-770 <VIT2>
A:Accession: C30320
A>Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 606-770 <VIT3>
R:Zain, S.B.; Salim, M.; Chou, W.G.; Sajdel-Sulkowska, E.M.; Majocha, R.E.; Marotta,
Proc. Natl. Acad. Sci. U.S.A. 85, 929-933, 1988
A:Title: Molecular cloning of amyloid cDNA derived from mRNA of the Alzheimer disease
A:Reference number: A31087; MUID:88124954; PMID:2893379

A:Accession: A31087
A:Molecule type: mRNA
A:Residues: 507-770 <A>
A:Cross-references: GB:M18734; NID:g178572; PIDN:AAAS1726.1; PID:g178573
A>Note: the authors translated the codon GAA for residue 599 as Gly, ACC for residue 603
8 as Val, GTG for residue 609 as Asn, AAT for residue 610 as Gly, and GGT for residue 65
A>Note: the cited Genbank accession number, J03594, is not in release 101.0
R:Masters, C.L.; Multhaup, G.; Simms, G.; Pottgiesser, J.; Martins, R.N.; Beyreuther, K.

Query Match 100.0%; Score 4058; DB 1; Length 770;
Best Local Similarity 100.0%; Pred. No. 5.4e-210;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MLPGLALLLLAAWTAARALEVPTDGNAGLLAEPTQAMFCGRLNMHMNVQNGKWDSPSGTK 60
Db 1 MLPGLALLLLAAWTAARALEVPTDGNAGLLAEPTQAMFCGRLNMHMNVQNGKWDSPSGTK 60
QY 61 TCIDTKEGILQYCOEYVPELQITNNVEANQPVTIQNWKGRKQCKTHPHFVPIYRCLVG 120
Db 61 TCIDTKEGILQYCOEYVPELQITNNVEANQPVTIQNWKGRKQCKTHPHFVPIYRCLVG 120
QY 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHTVAKETCEKSTNLDHYGMLLPCGIDKFR 180
Db 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHTVAKETCEKSTNLDHYGMLLPCGIDKFR 180
QY 181 GVEFVCCPLAEESDNVDSADAEEDSDVMWGGADTDYADGSEDKVVEVAEEVEAEVEE 240
Db 181 GVEFVCCPLAEESDNVDSADAEEDSDVMWGGADTDYADGSEDKVVEVAEEVEAEVEE 240
QY 241 EADDDDEDDGDEVEEAEPYEATERTTSTATTTTTTSTESVEEVVREVCSQAETGPC 300
Db 241 EADDDDEDDGDEVEEAEPYEATERTTSTATTTTTTSTESVEEVVREVCSQAETGPC 300
QY 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNFDTEEYCMVCGSAMQSLLKTTQEP LARD 360
Db 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNFDTEEYCMVCGSAMQSLLKTTQEP LARD 360
QY 361 PVKLPTTAASTPDADVCKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAEERQA 420
Db 361 PVKLPTTAASTPDADVCKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAEERQA 420
QY 421 KNLPAKADKAVIQHFOEKVESLEQEAANERQOLVETHMARVEAMLNDRRLALENYITAL 480
Db 421 KNLPAKADKAVIQHFOEKVESLEQEAANERQOLVETHMARVEAMLNDRRLALENYITAL 480
QY 481 QAVPPRPRHVFNMKKYVRAEQKDRQHTLKHFEHVRWVDPKAAQIRSQVMTHLRVIER 540
Db 481 QAVPPRPRHVFNMKKYVRAEQKDRQHTLKHFEHVRWVDPKAAQIRSQVMTHLRVIER 540
QY 541 MNQSLLYNPVPAVEEIQDEVELLQKEQNYSDVLANMISEPRISYGNALMPSLTET 600
Db 541 MNQSLLYNPVPAVEEIQDEVELLQKEQNYSDVLANMISEPRISYGNALMPSLTET 600
QY 601 KTTVELLPVNGEFLDLPWHSFGADSVPAANTENEVEVPDARPAADRLTTRPGSGLTN 660
Db 601 KTTVELLPVNGEFLDLPWHSFGADSVPAANTENEVEVPDARPAADRLTTRPGSGLTN 660
QY 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVGWVIATVIVITL 720
Db 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVGWVIATVIVITL 720
QY 721 VMLKKQYTSIHGGVVEVDAVTPPERHLSKMQQNGYENPTYKFFEQMQN 770
Db 721 VMLKKQYTSIHGGVVEVDAVTPPERHLSKMQQNGYENPTYKFFEQMQN 770

RESULT 2
A49795
Alzheimer's disease amyloid beta protein precursor - crab-eating macaque
C:Species: Macaca fascicularis (crab-eating macaque)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A49795
R:Podlisky, M.B.; Tolan, D.R.; Selkoe, D.J.

Am. J. Pathol. 138, 1423-1435, 1991
A:Title: Homology of the amyloid beta protein precursor in monkey and human supports
A:Reference number: A49795; MUID:91273117; PMID:1905108
A:Accession: A49795.
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-695 <POD>
A:Cross-references: GB:M58727; NID:g342062; PIDN:AAA36829.1; PID:g342063
C:Superfamily: Alzheimer's disease amyloid beta protein; animal kunitz-type proteinase
C:Keywords: alternative splicing

Query Match 88.5%; Score 3590.5; DB 1; Length 695;
Best Local Similarity 90.1%; Pred. No. 5.3e-185;
Matches 694; Conservative 1; Mismatches 0; Indels 75; Gaps 1;
QY 1 MLPGLALLLLAAWTAARALEVPTDGNAGLLAEPTQAMFCGRLNMHMNVQNGKWDSPSGTK 60
Db 1 MLPGLALLLLAAWTAARALEVPTDGNAGLLAEPTQAMFCGRLNMHMNVQNGKWDSPSGTK 60
QY 61 TCIDTKEGILQYCOEYVPELQITNNVEANQPVTIQNWKGRKQCKTHPHFVPIYRCLVG 120
Db 61 TCIDTKEGILQYCOEYVPELQITNNVEANQPVTIQNWKGRKQCKTHPHFVPIYRCLVG 120
QY 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHTVAKETCEKSTNLDHYGMLLPCGIDKFR 180
Db 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHTVAKETCEKSTNLDHYGMLLPCGIDKFR 180
QY 181 GVEFVCCPLAEESDNVDSADAEEDSDVMWGGADTDYADGSEDKVVEVAEEVEAEVEE 240
Db 181 GVEFVCCPLAEESDNVDSADAEEDSDVMWGGADTDYADGSEDKVVEVAEEVEAEVEE 240
QY 241 EADDDDEDDGDEVEEAEPYEATERTTSTATTTTTTSTESVEEVVREVCSQAETGPC 300
Db 241 EADDDDEDDGDEVEEAEPYEATERTTSTATTTTTTSTESVEEVVREVCSQAETGPC 300
QY 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNFDTEYCMVCGSAMQSLLKTTQEP LARD 360
Db 289 ----- 288
QY 361 PVKLPTTAASTPDADVCKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAEERQA 420
Db 289 ---VPTTAASTPDADVCKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAEERQA 345
QY 421 KNLPAKADKAVIQHFOEKVESLEQEAANERQOLVETHMARVEAMLNDRRLALENYITAL 480
Db 346 KNLPAKADKAVIQHFOEKVESLEQEAANERQOLVETHMARVEAMLNDRRLALENYITAL 405
QY 481 QAVPPRPRHVFNMKKYVRAEQKDRQHTLKHFEHVRWVDPKAAQIRSQVMTHLRVIER 540
Db 406 QAVPPRPRHVFNMKKYVRAEQKDRQHTLKHFEHVRWVDPKAAQIRSQVMTHLRVIER 465
QY 541 MNQSLLYNPVPAVEEIQDEVELLQKEQNYSDVLANMISEPRISYGNALMPSLTET 600
Db 466 MNQSLLYNPVPAVEEIQDEVELLQKEQNYSDVLANMISEPRISYGNALMPSLTET 525
QY 601 KTTVELLPVNGEFLDLPWHSFGADSVPAANTENEVEVPDARPAADRLTTRPGSGLTN 660
Db 526 KTTVELLPVNGEFLDLPWHSFGADSVPAANTENEVEVPDARPAADRLTTRPGSGLTN 585
QY 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVGWVIATVIVITL 720
Db 586 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVGWVIATVIVITL 645
QY 721 VMLKKQYTSIHGGVVEVDAVTPPERHLSKMQQNGYENPTYKFFEQMQN 770
Db 646 VMLKKQYTSIHGGVVEVDAVTPPERHLSKMQQNGYENPTYKFFEQMQN 695

RESULT 3
S00550
Alzheimer's disease amyloid beta protein precursor - rat
N:Alternate names: beta-A4 amyloid protein
C:Species: Rattus norvegicus (Norway rat)

C:Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 13-Aug-1999
 C:Accession: S00550; A1245; A39820; S46251
 R:Shivers, B.D.; Hilbich, C.; Multhaup, G.; Salbaum, M.; Beyreuther, K.; Seeburg, P.H.
 EMBO J. 7, 1365-1370, 1988
 A:Title: Alzheimer's disease amyloidogenic glycoprotein: expression pattern in rat brain
 A:Reference number: S00550; MUID:88312583; PMID:2900758
 A:Accession: S00550
 A:Molecule type: mRNA
 A:Residues: 1-695 <SHI>
 A:CROSS-references: EMBL:X07648; NID:g55616; PIDN:CAA30488.1; PID:g55617
 R:Schubert, D.; Schroeder, R.; LeCorbriere, M.; Saitoh, T.; Cole, G.
 Science 241, 223-226, 1988
 A:Title: Amyloid beta protein precursor is possibly a heparan sulfate proteoglycan core
 A:Reference number: A1245; MUID:88264430; PMID:2968652
 A:Accession: A1245
 A:Molecule type: protein
 A:Residues: 18-37, 'X', 39-40, 'X', 42-44 <SCH>
 A:Note: evidence for heparan sulfate attachment
 R:Hesse, L.; Behr, D.; Masters, C.L.; Multhaup, G.
 FEBS Lett. 349, 109-116, 1994
 A:Title: The beta-A4 amyloid precursor protein binding to copper.
 A:Reference number: S46251; MUID:94320627; PMID:7913895
 A:Contents: annotation; copper binding sites
 A:Note: rat peptides were isolated but not sequenced
 R:Potempska, A.; Styles, J.; Mehta, P.; Kim, K.S.; Miller, D.L.
 J. Biol. Chem. 266, 8464-8469, 1991
 A:Title: Purification and tissue level of the beta-amyloid peptide precursor of rat brain
 A:Reference number: A39820; MUID:91217087; PMID:1673681
 A:Accession: A39820
 A>Status: preliminary
 A:Molecule type: protein
 A:Residues: 18-32 <POT>
 A:Experimental source: brain
 C:Comment: Deposition of amyloid protein as neurofibrillary tangles and/or plaques is characteristic of Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinase inhibitor; alternative splicing; amyloid; glycoprotein; transmembrane protein
 F:625-Domain: transmembrane #status predicted <TM>

Query Match 86.1%; Score 3493.5; DB 2; Length 695;
 Best Local Similarity 87.7%; Pred. No. 8.3e-180;
 Matches 675; Conservative 8; Mismatches 12; Indels 75; Gaps 1;

QY 1 MLPGLALLLAAMTARALEVPTDGNAGLLAEPOIAMFCGRNLNMHNVQNGKWDSPSGTK 60
 DB 1 MLPGLALLLAAMTARALEVPTDGNAGLLAEPOIAMFCGRNLNMHNVQNGKWDSPSGTK 60
 QY 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTIONWCKRGRKQCKTHPHFVTPYRCLVG 120
 DB 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTIONWCKRGRKQCKTHPHFVTPYRCLVG 120
 QY 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHVTAKETCSEKSTNLHDYGMLLPCGIDKFR 180
 DB 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHVTAKETCSEKSTNLHDYGMLLPCGIDKFR 180
 QY 181 GVEFVCCPLAESDSDAEDSDSDVWNGGADTDYADGSEDKVEVEAEVEAEVEE 240
 DB 181 GVEFVCCPLAESDSDAEDSDSDVWNGGADTDYADGSEDKVEVEAEVEAEVEE 240
 QY 241 EADDEDEDEGVEEAEPEYERATRTTSIATTTTTTTEVEEVEVREVCSEQAETGPC 300
 DB 241 EADDEDEDEGVEEAEPEYERATRTTSIATTTTTTTEVEEVEVREVCSEQAETGPC 300
 QY 301 RAMISRWFVTEGKCAFFYGGCGNRNFDTEYCMAYCGSAMSQLTKTTQEP LARD 360
 DB 301 RAMISRWFVTEGKCAFFYGGCGNRNFDTEYCMAYCGSAMSQLTKTTQEP LARD 360
 QY 361 PVKLPTTAATPDADYKYLETGPDENHAHFQAKERLEAKHRMSQVWREAEERQA 420
 DB 361 PVKLPTTAATPDADYKYLETGPDENHAHFQAKERLEAKHRMSQVWREAEERQA 420
 QY 289 ---VPTTAATPDADYKYLETGPDENHAHFQAKERLEAKHRMSQVWREAEERQA 345
 DB 289 ---VPTTAATPDADYKYLETGPDENHAHFQAKERLEAKHRMSQVWREAEERQA 345
 QY 421 KNLPAKDKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
 DB 421 KNLPAKDKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
 QY 346 KNLPAKDKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 405
 DB 346 KNLPAKDKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 405

QY 481 QAVPPRPHEVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 540
 DB 481 QAVPPRPHEVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 540
 QY 406 QAVPPRPHEVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 465
 DB 406 QAVPPRPHEVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 465
 QY 541 MNQSLSLLYNPAVAEETQDEVELLQEQNSDDVLANMISEPRISYGNDAIMPSTLET 600
 DB 541 MNQSLSLLYNPAVAEETQDEVELLQEQNSDDVLANMISEPRISYGNDAIMPSTLET 600
 QY 466 MNQSLSLLYNPAVAEETQDEVELLQEQNSDDVLANMISEPRISYGNDAIMPSTLET 525
 DB 466 MNQSLSLLYNPAVAEETQDEVELLQEQNSDDVLANMISEPRISYGNDAIMPSTLET 525
 QY 601 KTTVELLPVNGEFSDDDLQPMHSGADSVPAANTEVEPVDARPAADRGLTRPGSLTN 660
 DB 601 KTTVELLPVNGEFSDDDLQPMHSGADSVPAANTEVEPVDARPAADRGLTRPGSLTN 660
 QY 526 KTTVELLPVNGEFSDDDLQPMHSGADSVPAANTEVEPVDARPAADRGLTRPGSLTN 585
 DB 526 KTTVELLPVNGEFSDDDLQPMHSGADSVPAANTEVEPVDARPAADRGLTRPGSLTN 585
 QY 661 IKTEISEYKMDAEFRHSDSGVEYHHRQKLVFAEDVGSNGKGAIIGLMVGGVVIATVITL 720
 DB 661 IKTEISEYKMDAEFRHSDSGVEYHHRQKLVFAEDVGSNGKGAIIGLMVGGVVIATVITL 720
 QY 586 IKTEISEYKMDAEFRHSDSGVEYHHRQKLVFAEDVGSNGKGAIIGLMVGGVVIATVITL 645
 DB 586 IKTEISEYKMDAEFRHSDSGVEYHHRQKLVFAEDVGSNGKGAIIGLMVGGVVIATVITL 645
 QY 721 VMLKKQYTSIHGVEYVDAVTPPEERHLSKMQQNGYENPTYKFFEQMQN 770
 DB 721 VMLKKQYTSIHGVEYVDAVTPPEERHLSKMQQNGYENPTYKFFEQMQN 770
 QY 646 VMLKKQYTSIHGVEYVDAVTPPEERHLSKMQQNGYENPTYKFFEQMQN 695
 DB 646 VMLKKQYTSIHGVEYVDAVTPPEERHLSKMQQNGYENPTYKFFEQMQN 695

RESULT 4
 A27485
 Alzheimer's disease amyloid beta/A4 protein homolog precursor - mouse
 N:Alternate names: proteinase nexin II
 C:Species: Mus musculus (house mouse)
 C:Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 13-Aug-1999
 C:Accession: A27485; S19727; I49485
 R:Izumi, R.; Yamada, T.; Yoshikai, S.; Sasaki, M.; Hattori, M.; Sasaki, Y.
 Biochem. Biophys. Res. Commun. 149, 665-671, 1987
 A:Title: Complementary DNA for the mouse homolog of the human amyloid beta protein precursor
 A:Reference number: A27485; MUID:88106489; PMID:3322280
 A:Accession: A27485
 A:Molecule type: mRNA
 A:Residues: 1-695 <YAM>
 A:CROSS-references: GB:M18373; NID:g191568; PIDN:AAA37139.1; PID:g309085
 A:Experimental source: brain
 R:de Strooper, B.; van Leeuwen, F.; van den Berghe, H.
 Blochim. Biophys. Acta 1129, 141-143, 1991
 A:Title: The amyloid beta protein precursor or proteinase nexin II from mouse is closely related to the amyloid beta protein precursor of the mouse homolog of the human amyloid beta protein precursor
 A:Reference number: S19727; MUID:92096458; PMID:1756177
 A:Accession: S19727
 A:Molecule type: mRNA
 A:Residues: 1-210, 'G', 212-220, 'S', 222-396, 'A', 398-402, 'T', 404-448, 'A', 450-695 <STR>
 A:CROSS-references: EMBL:X59379
 R:Izumi, R.; Yamada, T.; Yoshikai, S.; Sasaki, M.; Hattori, M.; Sasaki, Y.
 Gene 112, 189-195, 1992
 A:Title: Positive and negative regulatory elements for the expression of the Alzheimer's disease amyloid beta protein precursor
 A:Reference number: I49485; MUID:92209998; PMID:1555768
 A:Accession: I49485
 A>Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-19 <RES>
 A:CROSS-references: GB:D10603; NID:g220328; PIDN:BAA01456.1; PID:g220329
 C:Genetics:
 A:Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinase
 C:Keywords: alternative splicing; amyloid; transmembrane protein

Query Match 85.5%; Score 3468.5; DB 2; Length 695;
 Best Local Similarity 87.3%; Pred. No. 1.8e-178;
 Matches 672; Conservative 6; Mismatches 17; Indels 75; Gaps 1;

QY 1 MLPGLALLLAAMTARALEVPTDGNAGLLAEPOIAMFCGRNLNMHNVQNGKWDSPSGTK 60
 DB 1 MLPGLALLLAAMTARALEVPTDGNAGLLAEPOIAMFCGRNLNMHNVQNGKWDSPSGTK 60
 QY 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTIONWCKRGRKQCKTHPHFVTPYRCLVG 120
 DB 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTIONWCKRGRKQCKTHPHFVTPYRCLVG 120
 QY 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHVTAKETCSEKSTNLHDYGMLLPCGIDKFR 180
 DB 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHVTAKETCSEKSTNLHDYGMLLPCGIDKFR 180

A:Description: Complete nucleotide ad deduced amino acid sequence of rat amyloid precursor protein
A:Reference number: S42880
A:Accession: S42880
A:Molecule type: mRNA
A:Residues: 1-765 <SAN>
A:Cross-references: EMBL:X77934
R:Sandbrink, R.; Masters, C.L.; Beyreuther, K.
Biochim. Biophys. Acta 1219, 167-170, 1994
A:Title: Complete nucleotide and deduced amino acid sequence of rat amyloid protein precursor
A:Reference number: S47528; MUID:94368849; PMID:8086458
A:Accession: S47528
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-765 <SA2>
A:Cross-references: EMBL:X77934
C:Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinase
C:Keywords: alternative splicing
F:312-362/Domain: animal Kunitz-type proteinase inhibitor homology <BPI>

Query Match 48.6%; Score 1971.5; DB 2; Length 765;
Best Local Similarity 49.7%; Pred. No. 3.3e-98;
Matches 401; Conservative 134; Mismatches 173; Indels 99; Gaps 20;

QY 5 LALLLLAAWTAARALEV-----PTDGNAG--LLAEPOIAMFCGLNMHMVQNGKWDSDP 56
DB 15 LVVLLLLGLTAPAAALAGYTEALANAGTGFVAEPOIAMFCGLKNHNVNIQTCKWEPPD 74
QY 57 SGTKCTCDTKEGILQYCOEYVPELQINNVYEAOPVTIQNWKCRGRKQCKTHPHFVTPYR 116
DB 75 TGTKSLCGLTKEEVLYCOEYVPELQINNVYEAOPVNIQNSWCRDRDKQCRS--HIVTPFK 132
QY 117 CLVGEFVSDDLVPDKCKFLHQERMVOCETHLHHVTHVAKETCSKSTNLHDYGMLLPCGI 176
DB 133 CLVGEFVSDDLVPDKCKFLHQERMVOCETHLHHVTHVAKETCSKSTNLHDYGMLLPCGI 192
QY 177 DKRGVEFVCCPLAE--ESDNVDSADAEDSDSVWVGADTDYA-DGSEDKVYVEVAEEEE 233
DB 193 DQFHGTIVYCPQKVVVDSSTWSKKEEEEE--DEEDYALDKSEFPTADLEDFDFT 248
QY 234 VAEVEEAEADDEDEDGSDVEEAEPYEE-----ATERTTSIATTTTTTSTESVEEV 287
DB 249 EAAADEDEDEEEEGVEEDRDYVYDFKDDYNEENPTPSSDGTISDKETAHDV 308
QY 288 REVCSQAEQTPCRAMTSRWYFDVTEGKCAFFYGGCGGNNRNFDTTEYCMVACVGSAMSQ 347
DB 309 KAVCSQAEQTPCRAMTSRWYFDVTEGKCAFFYGGCGGNNRNFDTTEYCMVACVGSAMSQ 366
QY 348 SLTKTTOEPLARQVPLKPTTAASTPDADVKYLETPTGDENEHAHFQAKERLEAKHHERMS 407
DB 367 -----PPTPLPT-----NDVDYFETSADNDNEHAFQAKERLEAKHHERMS 408
QY 408 QVREWEAEARQAKNLPKADKAVIQHFQKVESLEOEAAANRQQLVETHMARVEAMLND 467
DB 409 RVKKEWEAEALQAKNLPKAEQRTLIQHFQAVKALEKEAAEKQQLVETHLARVEAMLND 468
QY 468 RRLALENYITAOQVPPRPRHVFNMMLKVKYVRAEQDKRQHTLKHFEHVRMVDPKKAAQIR 527
DB 469 RRLALENYITAOQVPPRPRHVFNMMLKVKYVRAEQDKRQHTLKHFEHVRMVDPKKAAQIR 528
QY 528 SOVMTLRLVIERMNSLSLLYNVPVAAEIQDEVDLQKEQYSDVLLANMISPRIS 587
DB 529 SOVMTLRLVIERMNSLSLLYNVPVAAEIQDEVDLQKEQYSDVLLANMISPRIS 574
QY 588 YGNDALMPSLTETKTTVELLPVNGEISLDDLPQWHSFGADSVFANTENEPEVPDARPAAD 647
DB 575 ---DQFTSSISENPVDYR---VSSEES-EIPIPFHPF--HPFFSLSENE-----DTQPELY 621
QY 648 RGLTTRPGSLTN-----IKTEE---ISEVKMDAEFRHDSGYEVHHQKLVFFAEYDGS-- 697
DB 622 HPM--KKGSCMAEQDGLGIAEEKVINSKMKMDENVVIDETLDV--KEMIFNAERVGGLE 677
QY 698 -----NKGAIIGLVGGVGVATVIVITVLMLKKKQYTSIHGVEVVEDAAV 742

```

Qy 643 RPAADRGLTTPSGGLTNIKTEI--SEVKMDAEFRHDSGYEVHHOKLVFAEDVGS----- 697
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 613 -----GSGMAEQDG-GLTGABEKVINSKNMDENNVDETLDV--KEMIFNAERVGGLBEE 665
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 698 -----NKGATIGLVGGVVVIATVITVLMLKKKQYTSIHGGVVEVDAAVTP 744
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 666 PESVGPIUREDVSUSSNALIGLLVIAVAITVVISLMLKKRQYGTISHGIVEVDPMLTP 725
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 745 EERHLSKMQQNGYENPTYKFEQM 769
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 726 EERHLNKMQHGYENPTYKLEQM 750
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 10
A46362
amyloid precursor-like protein - mouse
C:Species: Mus musculus (house mouse)
C:Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 24-Nov-1999
C:Accession: A46362
R:Masco, W.; Bupp, K.; Magendantz, M.; Gusella, J.F.; Tanzi, R.E.; Solomon,
Proc. Natl. Acad. Sci. U.S.A. 89, 10758-10762, 1992
A:Title: Identification of a mouse brain cDNA that encodes a protein related
A:Reference number: A46362; MUID:93066322; PMID:1279693
A:Accession: A46362
A>Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-653 <WAS>
A:Experimental source: brain
A:Note: sequence inconsistent with the nucleotide translation
A:Note: sequence extracted from NCBI backbone (NCBIN:118683, NCBITP:118684)
C:Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type
C:Keywords: transmembrane protein

```

Query Match	28.4%	Score 1150.5;	DB 2;	Length 653;
Best Local Similarity	35.4%;	Pred. No. 2.6e-54;		
Matches	274;	Conservative 120;	Mismatches 228;	Indels 153; Gaps 19;

Qy	1	MLPGIALLLLAAWTARA-LVEPTDGNAGLAEPOIAMFCGLNHHMNVONGKWDSDPSGT	59
Db	22	LIP-LSULLLRQAUVGNLAVGSPSAAEAPGAQVAGCGRLTLUHLRDLRTGRWEPDPQRS	80
Qy	60	KTCIDTKEGIIQQCYQYVELQINVVYEANOPVTIQNWCKRGKRQCKTHPHF-VIPYRCL	118
Db	81	RRCLLDPQRVLEYCRQYMPELHIAVFEQAQAIPMRWCWGTRSGCAHPHIEVVPHEL	140
Qy	119	VGERVSODALLVPDKCKFLHQBRMDVCETHLHWHTVAKETSEKSTNLHDYGMLPLPCGIDK	178
Db	141	PGEVFSEALLPVEGCRFLHQBRMDQCESSTRHQEAQEACSSQGILLIHGSMGLLPFGCSDR	200
Qy	179	FRGVEFFCCPLAESDNVDSDAEDSDVW-WGGADTDYADGSEDKVVEVAEEEEEVAEV	237
Db	201	FRGVEYYCCP--PPATPNPMSGMAAGDPSTRSPFLGGR---AEGGED-----EEVESF	248
Qy	238	EEEEADDDEDEDDEEEAEPEEPEATERTTSIATTTTTTESVEVREVVCSEQAET	297
Db	249	PQPYDDIFYVEPPQAEHEEEEEERAPPSSHTP-----	281
Qy	298	GPCRAMISRWYFDVTEGKCADPFYGGCGGNRNFTDTEYCMVCGSAMSQLTKTTQEP	357
Db	282	----VMVSR-----VT-----	288
Qy	358	ARDPVKLPTTAASTPDAVDKYLETPGDENENHAHFQAKAKERLEAKHRMSOVMBWEBAE	417
Db	289	---PTPRPT-----DGVDFYFGMPGEIGEHEGFLRAKMGLEERMQRQINEMREWAMAD	339
Qy	418	RQAKNLKPADKKVIQHFOEKVESLEQEAAENERQLVETHMARVEAMLNRRRLALENIY	477
Db	340	SQSNLKPADQALNEHFQSILQILEQVSGERQLVETHATRVALINDOPRAALSGFL	399
Qy	478	TALQAVPPRPRRHNMLUKYVVRASOKORQHTLKHFHEHVHMVDPKKAQAIQSQVWTHLV	537
Db	400	AALQGDPPQABRVLMARLYRLARABQKHQTHLRYHQHVAAVDPEKAQMRFQVQTHLQIV	459

Qy	538	YERNQSLSLYVNPVAAEETQDEVDLLAQEQNYSDVLANMISEPRISYGNDA	536
		: : : : : : : : :	
Db	460	EERNQSLGLDQNPFLAQELRPQIQELL-----LAELGPGSEL-----DASVPGS	505
Qy	597	LTEIKTTVELLPVNGEESLDDLPWHSGADSVPAANTENEVEPVDARPAADRG	656
		: : : : : : : : :	
Db	506	SSEDK-----GSLQ-----PESKDDPPVTLP-----KGSTQESS	536
Qy	657	GLTIKITEISEVKMDAEFRHDSGYEVHH---QKLVPFAEDVGSNKGAIGL	713
		: : : : : : : : :	
Db	537	SSGREKLTPLEQYEQKNASAPRFPFHSIDIQDELAPSGTGVSRALSGLLMAGGG	596
Qy	714	TVIVITVLML-KKQYTSIIHGVVYEDAAVTPPEERHLSKMQQNGYENPTYKFEQ	767
		: : : :	
Db	597	SLIVLSLLLRKKPKPYGTISHGVVYEDPMLTLEBQULREQLRHGYENPTYRFL	651
		: : : : : : : : :	
RESULT 11			
CDEI1404			
C:Species: Mus musculus (house mouse)			
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Feb-1997			
C:Accession: JC1404			
R:Vidal, F.; Biangy, A.; Rassoulzadegan, M.; Cuzin, F.			
Biochem. Biophys. Res. Commun. 189, 1336-1341, 1992			
A:Title: A murine sequence-specific DNA binding protein shows extensive local similar			
A:Reference number: JC1404; MUID:93129193; PMID:1482349			
A:Accession: JC1404			
A:Molecule type: mRNA			
A:Residues: 1-511 <VID>			
C:Comment: This protein plays an important role in the early development of the mouse			
C:Keywords: DNA binding; transmembrane protein			
Query Match	27.4%;	Score 1112.5; DB 2; Length 511;	
Best Local Similarity	43.0%;	Pred. No. 2.le-52;	
Matches	264; Conservative	87; Mismatches 136; Indels 127; Gaps 21;	
Qy	174	CGIDKRGVEVCPPLAEESDNVSDAEDSDVWGGADTDYADGSEDKVVEVAEE	233
		: : : : : : :	
Db	6	CGVDQFHGTGYCCP---QTKTVD-----DSTMSEK-----EEEE	37
Qy	234	VAVEEEEAADDEDGDEVEEAEPYEATERTTSIATTTTTTTSVEEVEVREVCSE	293
		: : : : : : : : : : : : : : :	
Db	38	-----EEDEEEDYDLDKSEFFTEADLDGFTB---AAADEEEDEEEGVVD	86
Qy	294	QAETGPCRAMISRMYFDVTEGKAPFFYGGCGGNRNFDEEYCMAYCGSAMSQLKTT	353
		: : : :	
Db	87	-----RDYYD-----PF-----KGDDYNEENPTE-----PSSEGTI--S	114
Qy	354	QELIARDPVKLPTTAASDPDAVDKYLTPGDENHAFQKAKERLEAKHRERSQVWRE	413
		: : : : : : : : : : : : : : : : : : :	
Db	115	DKEIVHD-VKVPPTPLPTND-VDYFFTSADNCHAFQKAEQLEIRHNRNDRVKKEW	172
Qy	414	EEAEQAKNLPKADKAVIOHFQEKVESLEQEAANERQQLVETHMARVEAMLDRRLAL	473
		: : : : : : : : : : : : : : : : : :	
Db	173	EEAEQAKNLPKTERQTLIHQFQAMVKALEKAASEKQQLVETHLARVEAMLDRRLAL	232
Qy	474	ENYITALQAVPPRPRHVFNMMLKKYVRAEQDKRQHTLKHFEHVRMVDPKAAQIRSQVMTH	533
		: : : : : : : : : : : : : : : : : :	
Db	233	ENYLAALQSDPPRPHRLQALRRYVRAENKDRLHTIRHYQHVLAVDPEKAAQMKSQVMTH	292
Qy	534	LRVIYERMQNSLLLYNVPVAAEIQDEVDLQEQNYSDVLANMISEPRISYGNDA	593
		: : : : : : : : : : : : : : : : : :	
Db	293	LHVTEERNQSLSLYKVPVYQAQIEQIEDELQEQR-----ADM-----DQF	335
Qy	594	MPSLTETKTTVELLPVNGEESLDDLPWHSGADSVPAANTENEVEPVDARPAADRG	653
		: : : : : : : : : : : : : : : : : :	
Db	336	TSSISENPVDRVVSSESE-EIPFPFLHPTF-----PULSENE-----GSGMAEQ	379
Qy	654	PGSLGTLNKTTEEI-SEVMQDAEFRHDSGYEVHHQKLVFFAEDVGS	697
		: : : : : : : : : : : : : : : : : :	
Db	380	DG-SLIGAEKVINSKNMKNMNYIDETLDV--KEMIFNAERYGGLGEEPEESVGPLREDF	436

QY 238 EE-EAADDDEDDGDEVEEAPEYEEA-TERTTSIAFTTTTTTTSVEEVVREVCSEA 295
Db 297 DSCEGDNTEEDGAGESAEVSWDGGGKVVSLKSDSSSPSPAPAPAPKPKS 356
QY 296 ETGPCRAMISRWFYDTEGKCAPFFYGGCGGNRNFDEEYCMAYCGSAMSQSLKTTQE 355
Db 357 ESVTSTPQLS-----ASAAFAAANGSGT-----GAGAPPSTAQPTSD 396
QY 356 PLAROPVKLPTTAASPDAVDKYLETPGDENEHAFQAKERLEAKHREMSQVWREWE 415
Db 397 P-----YTFHDFHYEQSYKVSQKRLSEHREKVTVMKMDSD 435
QY 416 AEFQAKLPAADKA-----VIOHFQKVESLEQEAANERQOLVETHMARVEAMLNDR 468
Db 436 LEEKYQDMRLADPKAQSFKQRTARFQTSVQALBEEGNAEKHQLAAMHQQRVLAHINQR 495
QY 469 RRLALENYITALQAVPRPRHFMMLKYYVRAEQDKRQHTLKHFEH-VRMVDP-----KKA 524
Db 496 KREAMTCYQALTEQPPNAHVEKCLQKLLRALHKLDRALAHYRHLNLSGGPGGLEAAA 555
QY 525 QIRSQVMTLHVIRYERMNQSLLYNVPAAVEI-----QDEV-----562
Db 556 SERPRTLERLIDIRAVNOSMTMLKYPPELSAKIAQLMNDYILALRSKDDIPGSSIGMSE 615
QY 563 -----DELLOKEQNYSDDLNLAN 579
Db 616 EAAGILDKYRVEIERKVAEKRLAEKQREKQRAAREKLEKRLLEAKKVDMLKS 675
QY 580 MISE-----PRISYGNDAIM-----PSUTETKTTVTELLPVNG 611
Db 676 QVAEQSQOPTOSTSQAOQOQOQESLPGKELGPDALVTAANPNLETTKS-----726
QY 612 EFLSDDLQPHSGADSVPAANTENEVEPDARPAADRLTTPRPGSLTNKITEISEVKM 671
Db 727 EKLSLTE-----YGEATVSTTKVTQVLPTVDDDAVQRAVEDVAAA-----VAHQEA 773
QY 672 DAERFDHSGVEVHQKLVF-----FAEDVGSNK-----GAIIGLMVGGVVIATVITVLML 723
Db 774 EPQVQFHTMDLHGHSFSLREFQAHAHAKEGRNVVFTLSFAGIALMAAVFVGVA 833
QY 724 KKQYTSIH-HGVVEVDAVTP-----EERHLKMQQNGYENPTYKFFPE 766
Db 834 KWRTSRPHAQGTIEVDQNTVTHPIVREEKIYPNNQINGYENPTYKYFE 883
RESULT 14
S38344
CDEI-binding protein - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 19-May-1994 #sequence_revision 26-May-1995 #text_change 03-May-1996
C:Accession: S38344
R:Hanes, J.; von der Kammer, H.; Kristiansson, G.I.; Scheit, K.H.
Biochim. Biophys. Acta 1216, 154-156, 1993
A:Title: The complete cDNA coding sequence for the mouse CDEI binding protein.
A:Reference number: S38344; MUID:94032480; PMID:8218408
A:Accession: S38344
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-246 <HAN>
A:Cross-references: EMBL:222592
C:Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinase 1

Query Match 17.4%; Score 706; DB 2; Length 246;
Best Local Similarity 51.5%; Pred. No. 5e-31;
Matches 136; Conservative 35; Mismatches 51; Indels 42; Gaps 7;
QY 5 LALLLLAATARALEV-----PTDGNAG-----LLAEPOIAMFCGRLLNNHMMVQKWDSDP 56
Db 15 LVLVLLGLTAPAAALAGYIEALAAAGTGFVAEPAQIAMLGCKLNNHNVNIQTGWEPDP 74
QY 57 SGTCTCIDTREGILOYCQEVPELOITNVVEANQPVTIQNWCKRGKCKOCTHBPVYR 116
Db 75 TGTYSCLGTKEVLQYCOETPELOITNVMEANQPVNIQNSWCRDRKQCKS--HIVIPFK 132

QY 117 CLVGEFYSDALLVPDKCKFLHOERMDVCETHLHWHTVAKETSEKSTNLHDYGMLLPCGI 176
Db 133 CLVGEFYSDVLLVPDNCQFFHQRMEVCEKQRHWHITLVKEACLTGLTLYSYGMLLPCGV 192
QY 177 DKFRGVFVCCPLAEESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEEBEVAE 236
Db 193 DQFHGTVEVCCP---QTKTVDS-----DSTMSKEEEE---222
QY 237 VEDEEADDDDED-DEGDGEVEEEAE 259
Db 223 -EEDEEEDDEEDYDLDKSEFPTEAD 245
RESULT 15
A32282
Alzheimer's disease amyloid beta protein precursor - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Aug-1989 #sequence_revision 17-Aug-1989 #text_change 13-Aug-1999
C:Accession: A32282
R:Yamada, T.; Sakaki, H.; Dohura, K.; Goto, I.; Sakaki, Y.
Biochem. Biophys. Res. Commun. 158, 906-912, 1989
A:Title: Structure and expression of the alternatively-spliced forms of mRNA for the
A:Reference number: A32282; MUID:89149813; PMID:2493250
A:Accession: A32282
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-100 <YAM>
A:Cross-references: GB:M24397; NID:g200350; PID:AAA39929, 1; PID:g200351
C:Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinase
C:Keywords: alternative splicing
F:11-61/Domain: animal Kunitz-type proteinase inhibitor homology <BPI>
Query Match 12.3%; Score 501; DB 2; Length 100;
Best Local Similarity 92.0%; Pred. No. 1.5e-20;
Matches 92; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
QY 281 ESVEEVVREVCSEAETGPCRAMISRWFYDTEGKCAPFFYGGCGGNRNFDTBEYCMAY 340
Db 1 ESVEEVVREVCSEAETGPCRAMISRWFYDTEGKCVFFYGGCGGNRNFDTBEYCMAY 60
QY 341 CGSAMSQSLKTTQPLARDPVKLPPTTAASTPDAVDKYLE 380
Db 61 CGSVSTQSLKTTSEPLPDPKLPPTTAASTPDAVDKYLE 100
Search completed: January 28, 2003, 14:19:59
Job time : 25.1912 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:03:18 ; Search time 11.7782 Seconds
(without alignments)
2711.515 Million cell updates/sec

Title: US-09-904-987-2

Perfect score: 4058

Sequence: 1 MLPGLALLLAATARALEV.....KMQQNGYENPTYKFFEQMQN 770

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query %	Score	Match	Length	ID	Description
1	4058	100.0	770	1	A4_HUMAN	P05067 homo sapien
2	3937	97.0	770	1	A4_RAT	P08592 rattus norv
3	3927	96.8	770	1	A4_MOUSE	P12023 mus musculu
4	3921.5	96.6	751	1	A4_SAISC	Q95241 saimiri sci
5	1981.5	48.8	763	1	APP2_HUMAN	Q06481 homo sapien
6	1971.5	48.6	765	1	APP2_RAT	P15943 rattus norv
7	1704.5	42.0	695	1	APP2_MOUSE	Q06335 mus musculu
8	1155.5	28.5	650	1	APPL_HUMAN	P31693 homo sapien
9	1150.5	28.4	653	1	APPL_MOUSE	Q03157 mus musculu
10	785	19.3	686	1	A4_CAEEL	Q10651 caenorhabdi
11	737.5	18.2	886	1	A4_DROME	P14599 drosophila
12	457	11.3	87	1	A4_MACFA	P53601 macaca fasc
13	403	9.9	76	1	A4_MACMU	P29216 macaca mula
14	292	7.2	59	1	A4_BOVIN	Q28053 bos taurus
15	288	7.1	58	1	A4_RABIT	Q28748 oryctolagus
16	288	7.1	58	1	A4_SHEEP	Q28757 ovis aries
17	287	7.1	58	1	A4_CANFA	Q28280 canis famil
18	283	7.0	57	1	A4_PIG	Q29023 sus scrofa
19	283	7.0	57	1	A4_URUMA	Q29149 ursus marit
20	194.5	4.8	3911	1	AKA9_HUMAN	Q99996 h a-kinase
21	186	4.6	252	1	SPT2_HUMAN	Q43278 homo sapien
22	186	4.6	993	1	SCPI_MOUSE	Q62209 mus musculu
23	176	4.3	55	1	ISH1_STOHE	P31713 stoichactis
24	175.5	4.3	579	1	G160_HUMAN	Q08378 homo sapien
25	175	4.3	302	1	TPPI_RAT	Q02445 rattus norv
26	174.5	4.3	252	1	SPT2_MOUSE	Q3wu03 mus musculu
27	174.5	4.3	513	1	SPT1_HUMAN	Q43278 homo sapien
28	169	4.2	1875	1	MLP1_YEAST	Q02455 saccharomyc
29	168	4.1	304	1	TPPI_HUMAN	P10646 homo sapien
30	166	4.1	55	1	ISH2_STOHE	P81129 stoichactis
31	164.5	4.1	507	1	SPT1_MOUSE	Q9F097 mus musculu
32	163.5	4.0	346	1	ANBP_MERUN	Q25577 meriones un
33	163.5	4.0	1130	1	YL17_CAEEL	Q11102 caenorhabdi

RESULT 1

ID	A4_HUMAN	STANDARD;	PRT;	770 AA.
AC	P05067; P09000; Q16011;			
DT	13-AUG-1987 (Rel. 05, Created)			
DT	01-NOV-1991 (Rel. 20, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Alzheimer's disease amyloid A4 protein precursor (Protease nexin-II)			
DE	(PN-II) (APPI) (Contains: Beta-amyloid protein (Beta-APP) (A-beta))			
GN	APP OR A4 OR CVAP OR ADL			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain;			
RX	MEDLINE=871144572; PubMed=2881207;			
RA	Kang J., Lemaire H.-G., Unterbeck A., Salbaum J.M., Masters C.L.,			
RA	Grzeschik K.-H., Multhaup G., Beyreuther K., Mueller-Hill B.;			
RT	"The precursor of Alzheimer's disease amyloid A4 protein resembles a			
RT	cell-surface receptor.";			
RL	Nature 325:733-736(1987).			
[2]				
RN	SEQUENCE FROM N.A.			
RP	MEDLINE=88122639; PubMed=2893289;			
RX	Ponte P., Gonzalez-Dewhitt P., Schilling J., Miller J., Hsu D.,			
RA	Greenberg B., Davis K., Wallace W., Lieberburg I., Fuller F.,			
RA	Cordell B.;			
RT	"A new A4 amyloid mRNA contains a domain homologous to serine			
RT	proteinase inhibitors.";			
RL	Nature 331:525-527(1988).			
[3]				
RN	SEQUENCE FROM N.A.			
RP	MEDLINE=89128427; PubMed=2783775;			
RX	Lemaire H.G., Salbaum J.M., Multhaup G., Kang J., Bayney R.M.,			
RA	Unterbeck A., Beyreuther K., Mueller-Hill B.;			
RT	"The PreA4(695) precursor protein of Alzheimer's disease A4 amyloid			
RT	is encoded by 16 exons.";			
RL	Nucleic Acids Res. 17:517-522(1989).			
[4]				
RN	SEQUENCE FROM N.A.			
RP	MEDLINE=97263807; PubMed=9108164;			
RX	Hattori M., Tsukahara F., Furuhashi H., Hirose M.,			
RA	Saito M., Tsukuni S., Sakaki Y.;			
RT	"A novel method for making nested deletions and its application for			
RT	sequencing of a 300 kb region of human APP locus.";			
RL	Nucleic Acids Res. 25:1802-1808(1997).			
[5]				
RN	SEQUENCE OF 286-345 AND 365-366 FROM N.A.			
RP	MEDLINE=88122640; PubMed=2893290;			
RX	Tanzi R.E., McClatchey A.I., Lamperti E.D., Villa-Komaroff L.,			
RA	Gusella J.F., Nave R.L.;			
RT	"Protease inhibitor domain encoded by an amyloid protein precursor			
RT	mRNA associated with Alzheimer's disease.";			
RL	Nature 331:528-530(1988).			

ALIGNMENTS

Q02388 homo sapien
P12111 homo sapien
P45379 homo sapien
Q28864 macaca mula
Q01042 herpesvirus
P15989 gallus gall
P31569 oenothera v
Q14203 homo sapien
P49223 homo sapien
P31568 oenothera p
P38996 saccharomyc
P82968 melichaea c

34 163 4.0 2944 1 CA17_HUMAN
35 163 4.0 3176 1 CA36_HUMAN
36 162.5 4.0 297 1 TRT2_HUMAN
37 162 4.0 304 1 TPPI_MACMU
38 162 4.0 407 1 IE68_HSVSA
39 162 4.0 3137 1 CA36_CHICK
40 161.5 4.0 630 1 YCF2_OENVI
41 161.5 4.0 1278 1 DYNA_HUMAN
42 161 4.0 64 1 SPT3_HUMAN
43 160.5 4.0 721 1 YCF2_OENPI
44 160 3.9 802 1 NAB3_YEAST
45 159 3.9 197 1 MCP1_MELCP

RN [16] SEQUENCE OF 287-367 FROM N.A.
 RP MEDLINE=88122641; PubMed=28933291;
 RX Kitaguchi N., Takahashi Y., Tokushima Y., Shiojiri S., Ito H.,
 RA "Novel precursor of Alzheimer's disease amyloid protein shows
 RT protease inhibitory activity.";
 RL Nature 331:530-532(1988).
 RN [17]
 RP SEQUENCE OF 284-289 AND 365-770 FROM N.A.
 RX MEDLINE=87231971; PubMed=3035574;
 RA Robakis N.K., Ramakrishna N., Wolfe G., Wisniewski H.M.;
 RT "Molecular cloning and characterization of a cDNA encoding the
 RL cerebrovascular and the neuritic plaque amyloid peptides.";
 RL Proc. Natl. Acad. Sci. U.S.A. 84:4190-4194(1987).
 RN [18]
 RP SEQUENCE OF 507-770 FROM N.A.
 RX MEDLINE=88124954; PubMed=2893379;
 RA Zain S.B., Salim M., Chou W.G., Sajdel-Sulkowska E.M., Majocha R.E.,
 RA Marotta C.A.;
 RT "Molecular cloning of amyloid cDNA derived from mRNA of the Alzheimer
 RL disease brain: coding and noncoding regions of the fetal precursor
 RT mRNA are expressed in the cortex.";
 RL Proc. Natl. Acad. Sci. U.S.A. 85:929-933(1988).
 RN [19]
 RP SEQUENCE OF 672-681.
 RX MEDLINE=88035004; PubMed=3312495;
 RA Partridge W.M., Vinters H.V., Yang J., Eisenberg J., Choi T.B.,
 RA Tourtellotte W.W., Huebner V., Shively J.E.;
 RT "Amyloid angiopathy of Alzheimer's disease: amino acid composition
 RL and partial sequence of a 4,200-dalton peptide isolated from cortical
 RT microvessels.";
 RL J. Neurochem. 49:1394-1401(1987).
 RN [10]
 RP SEQUENCE OF 739-770 FROM N.A.
 RX MEDLINE=90236318; PubMed=2110105;
 RA Yoshikai S.-I., Sasaki H., Doh-Ura K., Furuya H., Sakaki Y.;
 RT "Genomic organization of the human amyloid beta-protein precursor
 RL gene.";
 RL Gene 87:257-263(1990).
 RN [11]
 RP SEQUENCE OF 1-10 FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=89016647; PubMed=3140222;
 RA Schon E.A., Mita S., Sadlock J., Herbert J.;
 RT "A cDNA specifying the human amyloid beta precursor protein (ABPP)
 RL encodes a 95-kDa polypeptide.";
 RL Nucleic Acids Res. 16:9351-9351(1988).
 RN [12]
 RP SEQUENCE OF 18-50.
 RX MEDLINE=87250462; PubMed=3597385;
 RA van Nostrand W.E., Cunningham D.D.;
 RT "Purification of protease nexin II from human fibroblasts.";
 RL J. Biol. Chem. 262:8508-8514(1987).
 RN [13]
 RP IDENTITY OF APP WITH NEXIN-II.
 RX MEDLINE=89384866; PubMed=2506449;
 RA Oltersdorf T., Fritz L.C., Schenk D.B., Lieberburg I.,
 RA Johnson-Wood K.L., Beattie E.C., Ward P.J., Blacher R.W., Dovey H.F.,
 RA Sinha S.;
 RT "The secreted form of the Alzheimer's amyloid precursor protein with
 RL the Kunitz domain is protease nexin-II.";
 RL Nature 341:144-147(1989).
 RN [14]
 RP PROTEASE-SPECIFICITY OF INHIBITOR DOMAIN.
 RX MEDLINE=90211252; PubMed=1969731;
 RA Kido H., Fukutomi A., Schilling J., Wang Y., Cordell B., Katunuma N.;
 RT "Protease-specificity of Kunitz inhibitor domain of Alzheimer's
 RL disease amyloid protein precursor.";
 RL Biochem. Biophys. Res. Commun. 167:716-721(1990).
 RN [15]
 RP COMPLEX WITH G(O).
 RX MEDLINE=93188965; PubMed=8446172;
 RA Nishimoto-I., Okamoto T., Matsuura Y., Takahashi S., Okamoto T.,

RA Murayama Y., Ogata E.;
 RT "Alzheimer amyloid protein precursor complexes with brain GTP-binding
 RL protein G(O).";
 RL Nature 362:75-79(1993).
 RN [16]
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 28-133.
 RX MEDLINE=99215582; PubMed=10201399;
 RA Rossjohn J., Cappai R., Feil S.C., Henry A., McKinstry W.J.,
 RA Galatis D., Hesse L., Multhaup G., Beyreuther K., Masters C.L.,
 RA Parker M.W.;
 RT "Crystal structure of the N-terminal, growth factor-like domain of
 RL Alzheimer amyloid precursor protein.";
 RL Nat. Struct. Biol. 6:327-331(1999).
 RN [17]
 RP X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS) OF 287-344.
 RX MEDLINE=91104913; PubMed=2125487;
 RA Hynes T.R., Randal M., Kennedy L.A., Eigenbrot C., Kossiakof A.A.;
 RT "X-ray crystal structure of the protease inhibitor domain of
 RL Alzheimer's amyloid beta-protein precursor.";
 RL Biochemistry 29:10018-10022(1990).
 RN [18]
 RP STRUCTURE BY NMR OF 289-344.
 RX MEDLINE=92031488; PubMed=1718421;
 RA Heald S.L., Tilton R.F. Jr., Hammond L.S., Lee A., Bayney R.M.,
 RA Kamark M.E., Ramabhadran T.V., Dreyer R.N., Davis G., Unterbeck A.,
 RA Tamburini P.P.;
 RT "Sequential NMR resonance assignment and structure determination of
 RL the Kunitz-type inhibitor domain of the Alzheimer's beta-amyloid
 RL precursor protein.";
 RL Biochemistry 30:10467-10478(1991).
 RN [19]
 RP STRUCTURE BY NMR OF 672-699.
 RX MEDLINE=94281210; PubMed=7516706;
 RA Talafous J., Marcinkowski K.J., Klopman G., Zagorski M.G.;
 RT "Solution structure of residues 1-28 of the amyloid beta-peptide.";
 RL Biochemistry 33:7788-7796(1994).
 RN [20]
 RP STRUCTURE BY NMR OF 696-706.
 RX MEDLINE=97128622; PubMed=8973180;
 RA Kohno T., Kobayashi K., Maeda T., Sato K., Takashima A.;
 RT "Three-dimensional structures of the amyloid beta peptide (25-35) in
 RL membrane-mimicking environment.";
 RL Biochemistry 35:16094-16104(1996).
 RN [21]
 RP STRUCTURE BY NMR OF 672-711.
 RX MEDLINE=98359783; PubMed=9693002;
 RA Coles M., Bicknell W., Watson A.A., Fairlie D.P., Craik D.J.;
 RT "Solution structure of amyloid beta-peptide(1-40) in a water-micelle
 RL environment. Is the membrane-spanning domain where we think it is?";
 RL Biochemistry 37:11064-11077(1998).
 RN [22]
 RP STRUCTURE BY NMR OF 672-699.
 RX MEDLINE=20400066; PubMed=10940222;
 RA Poulsen S.-A., Watson A.A., Craik D.J.;
 RT "Solution structures in aqueous SDS micelles of two amyloid beta
 RL peptides of Abeta(1-28) mutated at the alpha-secretase cleavage
 RL site.";
 RL J. Struct. Biol. 130:142-152(2000).
 RN [23]
 RP STRUCTURE BY NMR OF 681-706.
 RX MEDLINE=20400065; PubMed=10940221;
 RA Zhang S., Iwata K., Lachenmann M.J., Peng J.W., Li S., Stimson E.R.,
 RA Lu Y., Felix A.M., Maggio J.E., Lee J.P.;
 RT "The Alzheimer's peptide a beta adopts a collapsed coil structure in
 RL water.";
 RL J. Struct. Biol. 130:130-141(2000).
 RN [24]
 RP SIGNAL SEQUENCE CLEAVAGE SITE, AND TOPOLOGY.
 RX MEDLINE=88296437; PubMed=2900137;
 RA Dyrks T., Weidemann A., Multhaup G., Salbaum J.M., Lemaire H.-G.,
 RA Kang J., Mueller-Hill B., Masters C.L., Beyreuther K.;
 RT "Identification, transmembrane orientation and biogenesis of the
 RL amyloid A4 precursor of Alzheimer's disease."

Query Match		100.0%;	Score 4058;	DB 1;	Length 770;							
Best Local Similarity		100.0%;	Pred. No. 3.1e-204;									
Matches	770;	Conservative	0;	Mismatches	0;							
				Indels	Gaps							
QY	1	MLPGLALLLLAAWTA	RALEVP	TGDNAG	LLAE	QPIAMFCGRLN	HMNMVQ	NGKWDSP	SGTK	60		
Db	1	MLPGLALLLLAAWTA	RALEVP	TGDNAG	LLAE	QPIAMFCGRLN	HMNMVQ	NGKWDSP	SGTK	60		
QY	61	TCIDTREGILQYCOE	YVPELQ	ITNV	VEANQ	PVTI	QIOWCK	RGRKCK	THPHFV	IPYRCLVG	120	
Db	61	TCIDTREGILQYCOE	YVPELQ	ITNV	VEANQ	PVTI	QIOWCK	RGRKCK	THPHFV	IPYRCLVG	120	
QY	121	EFVSDALLVPDKCF	LHQRMD	VCETHL	HWHTV	AKETCSEK	STNLHD	YGMLLP	CGCIDK	PR	180	
Db	121	EFVSDALLVPDKCF	LHQRMD	VCETHL	HWHTV	AKETCSEK	STNLHD	YGMLLP	CGCIDK	PR	180	
QY	181	GVFVCCPLAESD	NVDSADA	EDDS	VWVG	GADTDY	ADGSE	KVYVEA	EEVEA	VEEVE	240	
Db	181	GVFVCCPLAESD	NVDSADA	EDDS	VWVG	GADTDY	ADGSE	KVYVEA	EEVEA	VEEVE	240	
QY	241	EADDDDEDDGDE	VEEAPE	YEATERT	STATTTT	TESVEE	VVREVC	SEQAET	GPC	300		
Db	241	EADDDDEDDGDE	VEEAPE	YEATERT	STATTTT	TESVEE	VVREVC	SEQAET	GPC	300		
QY	301	RAMISRWYFDVTE	GKCAPF	YFGCGG	NRNNFDT	EYCM	AVCGS	SMSQSL	LLKTQ	QPELARD	360	
Db	301	RAMISRWYFDVTE	GKCAPF	YFGCGG	NRNNFDT	EYCM	AVCGS	SMSQSL	LLKTQ	QPELARD	360	
QY	361	PVKLPPTAAS	TPDAVKY	LETPG	DENEH	AFQAKER	LEAKH	IRMSQ	VYMRWE	EAERQA	420	
Db	361	PVKLPPTAAS	TPDAVKY	LETPG	DENEH	AFQAKER	LEAKH	IRMSQ	VYMRWE	EAERQA	420	
QY	421	KNLPKADKKA	VIQHFQ	KEKVES	LEQEA	NAERQ	QLVETH	MAV	EAMLDNR	RRLALENYIT	480	
Db	421	KNLPKADKKA	VIQHFQ	KEKVES	LEQEA	NAERQ	QLVETH	MAV	EAMLDNR	RRLALENYIT	480	
QY	481	QAVPPRPRHVN	NLKKYV	RAEQ	KDRQ	HTLKH	FEHVR	MDVP	KKAAQ	IRSQVMT	HLRVIYER	540
Db	481	QAVPPRPRHVN	NLKKYV	RAEQ	KDRQ	HTLKH	FEHVR	MDVP	KKAAQ	IRSQVMT	HLRVIYER	540
QY	541	MNQSLLYNV	PAVEI	QDEVD	ELLQ	KQNSD	VDVLAN	WIS	SEPRISY	NDALM	PSLTET	600
Db	541	MNQSLLYNV	PAVEI	QDEVD	ELLQ	KQNSD	VDVLAN	WIS	SEPRISY	NDALM	PSLTET	600
QY	601	KTTVELLPVNGE	SLDLP	QWHSF	GADSV	PANTENE	VEP	VDAR	PAAD	RGLTTR	PGSLTN	660
Db	601	KTTVELLPVNGE	SLDLP	QWHSF	GADSV	PANTENE	VEP	VDAR	PAAD	RGLTTR	PGSLTN	660
QY	661	IKTEEISEVK	MDAEFR	HDG	SYEVH	HQKLV	FFAED	VGS	NKGAI	IGLM	VGGVVIAT	720
Db	661	IKTEEISEVK	MDAEFR	HDG	SYEVH	HQKLV	FFAED	VGS	NKGAI	IGLM	VGGVVIAT	720
QY	721	VMLKKQYTSIH	HGVVE	DAVTP	PERHLS	KMQQNG	YENPTY	KYFF	EQMON	770		
Db	721	VMLKKQYTSIH	HGVVE	DAVTP	PERHLS	KMQQNG	YENPTY	KYFF	EQMON	770		
RESULT 2												
A4_RAT	A4_RAT	STANDARD;		PRT;	770	AA.						
AC	P08592;											
DT	01-AUG-1988	(Rel. 08, Created)										
DT	01-DEC-1992	(Rel. 24, Last sequence update)										
DT	16-OCT-2001	(Rel. 40, Last annotation update)										
DE	Alzheimer's disease amyloid A4 protein homolog precursor											
DE	(Amyloidogenic glycoprotein) (AG).											
GN	APP.											
OS	Rattus norvegicus (Rat).											
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;											
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.											
OX	NCBI_TaxID=10116;											
RN	[1]											

SEQUENCE OF 1-289 AND 365-770 FROM N.A.
TISSUE=Brain;
RX MEDLINE=88312583; PubMed=2900758;
RA Shivers B.D., Hilbich C., Mulchaup G., Salbaum J.M., Beyreuther K., Seeburg P.H.;
RT "Alzheimer's disease amyloidogenic glycoprotein: expression pattern in rat brain suggests a role in cell contact.";
RL EMBO J. 7:1365-1370(1988).
RN [2]
RP SEQUENCE OF 289-364 FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=89183625; PubMed=2648331;
RA Kang J., Mueller-Hill B.;
RT "The sequence of the two extra exons in rat preA4.";
RL Nucleic Acids Res. 17:2130-2130(1989).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- ALTERNATIVE PRODUCTS: 5 ISOFORMS; APP(395), APP(563), APP(695), APP(751) AND APP(770) (SHOWN HERE); ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- DOMAIN: THE CLATHRIN-BINDING SITE IS ESSENTIAL FOR ITS ASSOCIATION WITH X11-ALPHA, -BETA, AND -GAMMA. THE SEQUENCE SPECIFIC RECOGNITION EXTENDS TO PEPTIDE RESIDUES THAT ARE C-TERMINAL TO THE NPXY MOTIF. THIS INTERACTION APPEARS TO BE INDEPENDENT OF PHOSPHORYLATION (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE APP FAMILY.
CC -1- SIMILARITY: CONTAINS 1 BPTI/KUNITZ INHIBITOR DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed, usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
CC -----
DR EMBL; X07648; CAA30488.1; -;
DR EMBL; X14066; CAA32229.1; -;
DR PIR; S00550; S00550.
DR PIR; S03607; S03607.
DR HSSP; P05067; IAAIP.
DR InterPro; IPR001868; A4_APP.
DR InterPro; IPR001255; Beta-APP.
DR InterPro; IPR002223; Kunitz_BPTI.
DR Pfam; PF00014; Kunitz_Bpti; 1.
DR Pfam; PF02177; A4_EXTRA; 1.
DR Pfam; PF03494; Beta-APP; 1.
DR PRINTS; PR00203; AMYLOIDA4.
DR PRINTS; PR00759; BASICPTASE.
DR PRODOM; PD000222; Kunitz_BPTI; 1.
DR SMART; SM00006; A4_EXTRA; 1.
DR SMART; SM00131; KU; 1.
DR PROSITE; PS00319; A4_EXTRA; 1.
DR PROSITE; PS00320; A4_INTRA; 1.
DR PROSITE; PS00280; BPTI_KUNITZ_1; 1.
DR PROSITE; PS02079; BPTI_KUNITZ_2; 1.
KW Glycoprotein; Amyloid; Neurone; Transmembrane; Signal; Alternative splicing; Serine protease inhibitor.
FT SIGNAL 1 17
FT CHAIN 18 770
FT ALZHEIMER'S DISEASE AMYLOID A4 PROTEIN
FT HOMOLOG.
FT DOMAIN 18 699
FT TRANSMEM 700 723
FT DOMAIN 724 770
FT DOMAIN 673 715
FT DOMAIN 287 345
FT SITE 759 762
FT DISULFID 291 341
FT DISULFID 300 324
FT DISULFID 316 337
FT CARBOHYD 542 542
FT CARBOHYD 571 571
FT VARSPLIC 289 289
FT VARSPLIC 290 364
FT MISSING (IN ISOFORM APP(695)).

DR PIR; A27485; A27485.
 DR PIR; S04855; S04855.
 DR PIR; S19727; S19727.
 DR HSP; P05067; IAAP.
 DR MGD; MGI:88059; App.
 DR InterPro; IPR001868; A4_APP.
 DR InterPro; IPR001255; Beta-APP.
 DR InterPro; IPR002223; Kunitz_BPTI.
 DR Pfam; PF00014; Kunitz_BPTI.
 DR Pfam; PF02177; A4_EXTRA; 1.
 DR Pfam; PF03494; Beta-APP; 1.
 DR PRINTS; PR00203; AMYLOIDA4.
 DR PRINTS; PR00759; BASICPTASE.
 DR ProDom; PD000222; Kunitz_BPTI; 1.
 DR SMART; SM00006; A4_EXTRA; 1.
 DR SMART; SM00131; KU; 1.
 DR PROSITE; PS00319; A4_EXTRA; 1.
 DR PROSITE; PS00320; A4_INTRA; 1.
 DR PROSITE; PS00280; BPTI_KUNITZ_1; 1.
 DR PROSITE; PS00279; BPTI_KUNITZ_2; 1.
 KW Glycoprotein; Amyloid; Neurone; Transmembrane; Signal;
 KW Alternative splicing; Serine protease inhibitor.
 FT SIGNAL 1 17
 FT CHAIN 18 770
 FT
 FT DOMAIN 18 699
 FT EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 700 723
 FT DOMAIN 724 770
 FT DOMAIN 673 715
 FT DOMAIN 287 345
 FT SITE 759 762
 FT DISULFID 291 341
 FT DISULFID 300 324
 FT DISULFID 316 337
 FT CARBOHYD 542 542
 FT CARBOHYD 571 571
 FT VARSPLIC 289 289
 FT VARSPLIC 290 364
 FT VARSPLIC 346 380
 FT VARSPLIC 346 380
 SQ SEQUENCE 770 AA; 86752 MW; 26C50DE0890CAF7A CRC64;

Query Match 96.8%; Score 3927; DB 1; Length 770;
 Best Local Similarity 96.8%; Pred. No. 2.1e-197;
 Matches 745; Conservative 7; Mismatches 18; Indels 0; Gaps 0;

QY 1 MLPGLALLLAANTARALEVPTDGNAGLLAEPOIANFCGLNMMNVQNGKWDSPSGTK 60
 DB 1 MLPGLALLLAANTARALEVPTDGNAGLLAEPOIANFCGLNMMNVQNGKWDSPSGTK 60
 QY 61 TCIDTREGILQYCOEYVPELQINNVYEAQNPVTIQNCKRGRKCKTHPHFVTPYRCLVG 120
 DB 61 TCIDTREGILQYCOEYVPELQINNVYEAQNPVTIQNCKRGRKCKTHPHFVTPYRCLVG 120
 QY 121 EFVSDALLVPDKFLHQRMDYCYETHLHWHTVAKETSEKSTNLHDYGMLLPCGIDKFR 180
 DB 121 EFVSDALLVPDKFLHQRMDYCYETHLHWHTVAKETSEKSTNLHDYGMLLPCGIDKFR 180
 QY 181 GFVEVCPLAESDNDVSDAEDDDSDVWVGADTDYAGSEDKDYVEVAEEVEEVEE 240
 DB 181 GFVEVCPLAESDNDVSDAEDDDSDVWVGADTDYAGSEDKDYVEVAEEVEEVEEVEE 240
 QY 241 EADDEDEDGDEVEEAPEPEATRTTSIATTTTTTETTESVEEVVREVCSEAETGPC 300
 DB 241 EADDEDEDGDEVEEAPEPEATRTTSIATTTTTTETTESVEEVVREVCSEAETGPC 300
 QY 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNFDTEEYCMVCGSAMQSLLKTTQPELARD 360
 DB 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNFDTEEYCMVCGSAMQSLLKTTQPELARD 360
 QY 361 PVKLPTTAATPDVAVKYLETPGDENEHAHFQAKERLEAKHRMSQVMEWEAEERQA 420
 DB 361 PVKLPTTAATPDVAVKYLETPGDENEHAHFQAKERLEAKHRMSQVMEWEAEERQA 420

QY 421 KNLPRADKKAVIQHFQEKVESLEQEAANERQOLVETHMARVEAMLNDRRRRLALENYITAL 480
 DB 421 KNLPRADKKAVIQHFQEKVESLEQEAANERQOLVETHMARVEAMLNDRRRRLALENYITAL 480
 QY 481 QAVPPRPRHFNMLKYYRAEQDKRQHTLKHFHEHVMYVDPKKAQIRSQVMTHLRVIER 540
 DB 481 QAVPPRPRHFNMLKYYRAEQDKRQHTLKHFHEHVMYVDPKKAQIRSQVMTHLRVIER 540
 QY 541 MNQSLSLYNNPVAEETQDEVDLLOKEQYSDDVLANMISEPRISYGNALMPSLTET 600
 DB 541 MNQSLSLYNNPVAEETQDEVDLLOKEQYSDDVLANMISEPRISYGNALMPSLTET 600
 QY 601 KTTVELLPVNGEFSLDDLPWHSFGADSVPAANTENEVEPVDARPAADRGLTRPCSGLTN 660
 DB 601 KTTVELLPVNGEFSLDDLPWHSFGADSVPAANTENEVEPVDARPAADRGLTRPCSGLTN 660
 QY 661 IKTEEISEVKMDAEPHSDGYEVHHQKLVFFAEDVGSNGKGAIIGLMVGGVVIATVITL 720
 DB 661 IKTEEISEVKMDAEPHSDGYEVHHQKLVFFAEDVGSNGKGAIIGLMVGGVVIATVITL 720
 QY 721 VMLKKKQYTSIHGVEVVDAAVTPPEERHLSKMQNGYENPTYKPFQEQMN 770
 DB 721 VMLKKKQYTSIHGVEVVDAAVTPPEERHLSKMQNGYENPTYKPFQEQMN 770

RESULT 4
 A4_SAISC STANDARD; PRT; 751 AA.
 AC Q95241;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Alzheimer's disease amyloid A4 protein precursor [Contains: Beta-amyloid protein (Beta-APP) (A-beta)].
 GN APP.
 OS Saimiri sciureus (Common squirrel monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Saimiri.
 OX NCBI_TaxID=9521;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP TISSUE=Liver, and Kidney;
 RX MEDLINE=96108492; PubMed=8532114;
 RA Levy E., Amorim A., Frangione B., Walker L.C.;
 RT "Beta-amyloid precursor protein gene in squirrel monkeys with cerebral amyloid angiopathy";
 RL Neurobiol. Aging 16:805-808(1995).
 CC -!- FUNCTION: FUNCTIONAL NEURONAL RECEPTOR WHICH COUPLES TO INTRACELLULAR SIGNALING PATHWAY THROUGH THE GTP-BINDING PROTEIN G(O).
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- DOMAIN: THE CLATHRIN-BINDING SITE IS ESSENTIAL FOR ITS ASSOCIATION WITH X11-ALPHA, -BETA, AND -GAMMA. THE SEQUENCE SPECIFIC RECOGNITION EXTENDS TO PEPTIDE RESIDUES THAT ARE C-TERMINAL TO THE NPXY MOTIF. THIS INTERACTION APPEARS TO BE INDEPENDENT OF PHOSPHORYLATION (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE APP FAMILY.
 CC -!- SIMILARITY: CONTAINS 1 BPTI/KUNITZ INHIBITOR DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL; S81024; AAD14347.1; -.
 DR HSP; P05067; IAAP.
 DR InterPro; IPR001868; A4_APP.
 DR InterPro; IPR001255; Beta-APP.
 DR InterPro; IPR002223; Kunitz_BPTI.
 DR Pfam; PF00014; Kunitz_BPTI; 1.

DR Pfam: PF02177; A4_EXTRA; 1.
 DR Pfam: PF03494; Beta-APP; 1.
 DR PRINTS: PR00203; AMYLOIDA4.
 DR PRINTS: PR00759; BASICPASE.
 DR ProDom: PD000222; Kunitz_BPTI; 1.
 DR SMART: SM00006; A4_EXTRA; 1.
 DR SMART: SM00131; KU; 1.
 DR PROSITE: PS00319; A4_EXTRA; 1.
 DR PROSITE: PS00320; A4_INTRA; 1.
 DR PROSITE: PS00280; BPTI_KUNITZ_1; 1.
 DR PROSITE: PS00279; BPTI_KUNITZ_2; 1.
 KW Glycoprotein; Amyloid; Neurone; Transmembrane; Alternative splicing;
 KW Signal; Serine protease inhibitor.
 FT SIGNAL 1 17 BY SIMILARITY.
 FT CHAIN 18 751 A4 PROTEIN.
 FT CHAIN 653 695 BETA-AMYLOID PROTEIN (POTENTIAL).
 FT DOMAIN 18 680 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 681 704 POTENTIAL.
 FT DOMAIN 705 751 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 287 345 BPTI/KUNITZ INHIBITOR.
 FT SITE 740 743 CLATHRIN-BINDING (BY SIMILARITY).
 FT ACT_SITE 301 302 REACTIVE_BOND.
 FT DISULFID 291 341 BY SIMILARITY.
 FT DISULFID 300 324 BY SIMILARITY.
 FT DISULFID 316 337 BY SIMILARITY.
 FT CARBOHYD 523 523 N-LINKED (GLCNAC. . .) (PROBABLE).
 FT CARBOHYD 552 552 N-LINKED (GLCNAC. . .) (PROBABLE).
 SQ SEQUENCE 751 AA; 84893 MW; 6C3E431085569049 CRC64;

Query Match 96.6%; Score 3921.5; DB 1; Length 751;
 Best Local Similarity 96.9%; Pred. No. 3.8e-197;
 Matches 746; Conservative 2; Mismatches 3; Indels 19; Gaps 1;

QY 1 MLPGLALLLAAMTARALEVPTDGNAGLLAEPOIAMFCGRLNHNHNVQNGKWDSPSGTK 60
 DB 1 MLPGLALLLAAMTARALEVPTDGNAGLLAEPOIAMFCGRLNHNHNVQNGKWDSPSGTK 60
 QY 61 TCIDTREGILOVQCOEYPELQITNVVEANQPTVIQNWCKRGRKQCKTHPHFVPIYRCLVG 120
 DB 61 TCIDTREGILOVQCOEYPELQITNVVEANQPTVIQNWCKRGRKQCKTHPHFVPIYRCLVG 120
 QY 121 EFVSDALLVPDKCKFLHQRMVDCVTHLHWHVAVKCTSEKSTNLHDYGMLLPCGIDKFR 180
 DB 121 EFVSDALLVPDKCKFLHQRMVDCVTHLHWHVAVKCTSEKSTNLHDYGMLLPCGIDKFR 180
 QY 181 GVEFVCCPLAEESDNDVSDADAEEDSDVWVGADTDYADGSKVVEVAEEVEEVEE 240
 DB 181 GVEFVCCPLAEESDNDVSDADAEEDSDVWVGADTDYADGSKVVEVAEEVEEVEEVEE 240
 QY 241 EADDDDEDDGDEVEEAEEPEEATERTTSIATTTTTTESVEEYVVCSEQAETGPC 300
 DB 241 EADDDDEDDGDEVEEAEEPEEATERTTSIATTTTTTESVEEYVVCSEQAETGPC 300
 QY 301 RAMISRWYFDVTEGKCAPFFYCGCGNRRNFDTEECVMAVCSSQSLLKTKTQEP LARD 360
 DB 301 RAMISRWYFDVTEGKCAPFFYCGCGNRRNFDTEECVMAVCSSQSLLKTKTQEP LARD 360
 QY 361 PVKLPPTAASTPDVADVKYLETGCDENEHAFQKAKERLEAKHRMSQVMRWEAEAEQA 420
 DB 361 PVKLPPTAASTPDVADVKYLETGCDENEHAFQKAKERLEAKHRMSQVMRWEAEAEQA 420
 QY 421 KNLPKADKXAVIQHFOGKVESLEQEAANERQQLVETHMARVEAMLNDRRRRLALENYITAL 480
 DB 421 KNLPKADKXAVIQHFOGKVESLEQEAANERQQLVETHMARVEAMLNDRRRRLALENYITAL 480
 QY 481 QAVPPRRHVNMLKKYVRAEQKDRQHTLKHFEHVRWVDPKKAQIRSQVTHLRIYIER 540
 DB 481 QAVPPRRHVNMLKKYVRAEQKDRQHTLKHFEHVRWVDPKKAQIRSQVTHLRIYIER 540
 QY 541 MNGSLSLLYNPAVAEIEIQDEVDLLQKQNYSDVLANMISEPRISYGNDAIMP SLTET 600
 DB 541 MNGSLSLLYNPAVAEIEIQDEVDLLQKQNYSDVLANMISEPRISYGNDAIMP SLTET 600
 QY 552 MNGSLSLLYNPAVAEIEIQDEVDLLQKQNYSDVLANMISEPRISYGNDAIMP SLTET 581
 DB 552 MNGSLSLLYNPAVAEIEIQDEVDLLQKQNYSDVLANMISEPRISYGNDAIMP SLTET 581

QY 601 KTTVELLPVNGEFSLDDLPWHSGADSVDPANTENEVEPVDPADPAADRLTTRPGSLTN 660
 DB 582 KTTVELLPVNGEFSLDDLPWHSGADSVDPANTENEVEPVDPADPAADRLTTRPGSLTN 641
 QY 661 IKTEIIESEVKMDAEPFRHDSGYEVHHQKLVFFAEDVGSNGKGAIIGLMVGGVVATVITVL 720
 DB 642 IKTEIIESEVKMDAEPFRHDSGYEVHHQKLVFFAEDVGSNGKGAIIGLMVGGVVATVITVL 701
 QY 721 VMLKKQYTSIHGVEVVEVDAATPEERHLSKMQNGYENPTYKFEQMQN 770
 DB 702 VMLKKQYTSIHGVEVVEVDAATPEERHLSKMQNGYENPTYKFEQMQN 751

RESULT 5
 APP2_HUMAN STANDARD; PRT; 763 AA.
 ID APP2_HUMAN Q06481;
 AC Q06481;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Amyloid-like protein 2 precursor (Amyloid protein homolog) (APPH)
 DE (CDEI-box binding protein) (CDEBP).
 GN APLP2 OR APLP2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=93250009; PubMed=8485127;
 RA Sprecher C.A., Grant F.J., Grimm G., O'Hara P.J., Norris F.,
 RA Norris K., Foster D.C.;
 RT "Molecular cloning of the cDNA for a human amyloid precursor protein
 RT homolog: evidence for a multigene family.";
 RL Biochemistry 32:4481-4486(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Ovary;
 RX MEDLINE=95217334; PubMed=7702756;
 RA von der Kammer H., Hanes J., Klaudiny J., Scheit K.H.;
 RT "A human amyloid precursor-like protein is highly homologous to a
 RT mouse sequence-specific DNA-binding protein.";
 RL DNA Cell Biol. 13:1137-1143(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=94035131; PubMed=8220435;
 RA Masco W., Gurubhagavatula S., Paradis M., Romano D.M., Sisodia S.S.,
 RA Hyman B.T., Neve R.L., Tanzi R.E.;
 RT "Isolation and characterization of APLP2 encoding a homologue of the
 RT Alzheimer's associated amyloid beta protein precursor.";
 RL Nat. Genet. 5:95-99(1993)
 CC -1- FUNCTION: MAY PLAY A ROLE IN THE REGULATION OF HEMOSTASIS. THE
 CC SOLUBLE FORM MAY HAVE INHIBITORY PROPERTIES TOWARDS COAGULATION
 CC FACTORS. MAY INTERACT WITH CELLULAR G-PROTEIN SIGNALING PATHWAYS.
 CC MAY BIND TO THE DNA 5'-GTCACATG-3'(CDEI BOX).
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN AND NUCLEAR
 CC (POTENTIAL).
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST 3 ISOFORMS; 1 (SHOWN HERE), 2 AND
 CC 3; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: IN PLACENTA, BRAIN, HEART, LUNG, LIVER, KIDNEY
 CC AND ENDOTHELIAL TISSUES.
 CC -1- SIMILARITY: BELONGS TO THE APP FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 BPTI/KUNITZ INHIBITOR DOMAIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

```

CC EMBL; S60059; AAC60589.1; -
DR EMBL; L09209; AAA35526.1; -
DR EMBL; 22572; CAA80295.1; -
DR EMBL; L27631; AAC41701.1; -
DR HSP; P05067; IMWP.
DR Genew; HGNC:598; APLP2.
DR MIM; 104776; -
DR InterPro; IPR001868; A4_APP.
DR InterPro; IPR002223; Kunitz_BPTI.
DR Pfam; PF00014; Kunitz_BPTI.
DR Pfam; PF02177; A4_EXTRA; 1.
DR PRINTS; PS00203; AMYLOIDA4.
DR PRINTS; PS00759; BASICPTASE.
DR ProDom; PD000222; Kunitz_BPTI.1.
DR SMART; SM00006; A4_EXTRA; 1.
DR SMART; SM00131; KU; 1.
DR PROSITE; PS00319; A4_EXTRA; 1.
DR PROSITE; PS00320; A4_INTRA; 1.
DR PROSITE; PS00280; BPTI_KUNITZ_1; 1.
DR PROSITE; PS00279; BPTI_KUNITZ_2; 1.
KW Transmembrane; Signal; Alternative splicing; DNA-binding;
KW Nuclear protein; Serine protease inhibitor.
FT SIGNAL 1 29 POTENTIAL.
FT CHAIN 30 763 AMYLOID-LIKE PROTEIN 2.
FT DOMAIN 30 692 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 693 716 POTENTIAL.
FT DOMAIN 717 763 CYTOSOLASMIC (POTENTIAL).
FT DOMAIN 215 280 ASP/GLU-RICH (HIGHLY ACIDIC).
FT DOMAIN 306 364 BPTI/KUNITZ INHIBITOR.
FT DOMAIN 215 231 POLY-GLU.
FT ACT_SITE 320 321 REACTIVE BOND (BY SIMILARITY).
FT DISULFID 310 360 BY SIMILARITY.
FT DISULFID 319 343 BY SIMILARITY.
FT DISULFID 335 356 BY SIMILARITY.
FT VARSPPLIC 308 363 MISSING (IN ISOFORM 2).
FT VARSPPLIC 613 624 MISSING (IN ISOFORM 3).
FT CONFLICT 543 543 S -> I (IN REF. 1).
SQ SEQUENCE 763 AA; 86955 MW; CA3A7D6DDB8A28D0 CRC64;

Query Watch 48.8%; Score 1981.5; DB 1; Length 763;
Best Local Similarity 50.4%; Pred. No. 3.6e-96;
Matches 408; Conservative 127; Mismatches 168; Indels 107; Gaps 21;

QY 5 LALLLLAAMTARALEY-----PTDGNAG---LLAEPQIAAFGRLNHHNMNVONGKWDSP 56
DB 15 LLLLLLGLTAPALAGYIEALANAGTGFAVEPQIAAFCEKLNHHVNIQTGWEPDP 74
QY 57 SGTKTCTDKTGELQYCEYVPELQITNVVEANQPTVIONWCKRGKCKTTPHFVPIYR 116
DB 75 TGTKSCFETKEEVLYQCQEMYPELQITNVNEANQRYSIDNWCRRDKCKS--RFVTPPK 132
QY 117 CLVGEFVSALLIYDPKCKFLQERNDVCETHLHHWHTVAKETSEKSTNLDHGMLLPCGI 176
DB 133 CLVGEFVSVDLLYPERKQCFHKEKMEVCENHQHWHVTVKAECLTQGMTLISYGMLLPCGV 192
QY 177 DKRGVEFVCCPLAEESDNVSDADADDSDVYMGGADTDYADGSDKDVVEVAEEVEAE 236
DB 193 DQFHGTEYVCCPOTKIGSVSKEEEDDE-----EEEEDEEDYDVYKSEFPTEAD 245
QY 237 VEE--EEA--DDEDDDDGDEVEEAEE-----EPYEATERTSIATTTTTSYVE 284
DB 246 LEDTEAAVDEDDDEDEEVEEDRDYDYDTFKGDYNE--ENPTEPGSDGTMDSKEIT 303
QY 285 EVREYVCSQAENGPCRAMISRHYFDVTGKCAPFFYGGCGGNRNFDTEECVMCGSA 344
DB 304 HDVKAVCSQAEMTGPCRANWPRYFDLSKGVRFYIGCGGNRNFNFESEDCMAVCKAM 363
QY 345 MSQSLLKTTQELARPVPKLTPTTAASDPDAVDKYLETPGDENEHAHFQAKERLEAKHRE 404
DB 364 I-----PPTPLPT-----NDVOVYFTSADNHEHARFQAKSOLEIRHN 403
QY 405 RMSQVMREWEAEARQAKNLPKADKAVIQHFQEKVESLEQEAAANRQQLVETHMARVEAM 464

```

```

DB 404 RMDRVKKEWEAEALQAKNLPRAERQCTLQHFQAMVKALEKAASEKQOLVETHLARVEAM 463
QY 465 LNDRRRLALENYITALQAVPPRPVHFMNKKYVRAEQKDRQHTLKHFHEHVMVDPKKAA 524
DB 464 LNDRRRLALENYLAALQSDPPRPHRIQLALRRYVRAENKRLHTIRHYQHVLAVDPEKAA 523
QY 525 QIRSOVMTHLRVIERMNSLSLLYNVPAVAEETODEVDELLQKEQNSDDVLANMISEP 584
DB 524 QMKSQVMTHLRVIERMNSLSLLYKVPYVAQETQEEIDELLQEQR-----ADM----- 572
QY 585 RISYGNDAIMPSTETKTVELLPVNGEFSLDDLPQWHSFGADSVDPANTNEVEPVDARP 644
DB 573 -----DQFTASISETPVDVR---VSSEES-EEPPFPFPF--HPEPALPENE-----DTQP 616
QY 645 AADRLTTRPGSLTN-----IKTEE-----ISEVKMDAEFRHDSGYEVHHQKLVFAEDVG 696
DB 617 ELYHPM--KGGSGVGEQDGLIGAEKVINSKNVDENMVIDETLDV--KEMIFNAERVG 672
QY 697 S-----NKGAILGLMVGGVVIATVITLVMLKKKQYTSIHGHVVVEVD 739
DB 673 GLEERESVGPLREDFSLSSALIGLLVIAVAIVIVISLVMURKRGYGTISHGIVEVD 732
QY 740 AAVTPEERHLSKMQONGYENPTYKFFEQMQ 769
DB 733 PMLTPEERHLNKMQHNGYENPTYKYLEQMQ 762

RESULT 6
APP2_RAT
ID APP2_RAT STANDARD; PRT; 765 AA.
AC P15943;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Amyloid-like protein 2 precursor (Sperm membrane protein YWK-II).
GN APLP2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE OF 1-627 FROM N.A.
RC STRAIN=Wistar; TISSUE=Brain, and Heart;
RX MEDLINE=94336849; PubMed=8086458;
RA Sandbrink R., Masters C.L., Beyreuther K.;
RT "Complete nucleotide and deduced amino acid sequence of rat amyloid
protein precursor-like protein 2 (APLP2/APPH): two amino acids length
difference to human and murine homologues.";
RL Biochim. Biophys. Acta 1219:167-170(1994).
RN [2]
RP SEQUENCE OF 575-765 FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=90207205; PubMed=1690887;
RA Yan Y.C., Bai Y., Wang L.F., Miao S.Y., Koide S.S.;
RT "Characterization of cDNA encoding a human sperm membrane protein
related to A4 amyloid protein.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:2405-2408(1990).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- ALTERNATIVE PRODUCTS: 4 ISOFORMS; A (SHOWN HERE), B, C AND D;
CC -!- ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -!- SIMILARITY: BELONGS TO THE APP FAMILY.
CC -!- SIMILARITY: CONTAINS 1 BPTI/KUNITZ INHIBITOR DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-slb.ch/
CC or send an email to license@isb-slb.ch).
CC -----
DB EMBL; X77934; CAA54906.1; -.

```

DR EMBL; M31322; AAA42352.1; -.
 DR PIR; A35981; A35981.
 DR HSSP; P05067; 1MWP.
 DR InterPro; IPR001868; A4_APP.
 DR InterPro; IPR002223; Kunitz_BPTI.
 DR Pfam; PF00014; Kunitz_BPTI; 1.
 DR Pfam; PF02177; A4_EXTRA; 1.
 DR PRINTS; PR00203; AMYLOIDA4.
 DR PRINTS; PR00759; BASICPTASE.
 DR ProDom; PD000222; Kunitz_BPTI; 1.
 DR SMART; SK00006; A4_EXTRA; 1.
 DR SMART; SK00131; KU; 1.
 DR PROSITE; PS00319; A4_EXTRA; 1.
 DR PROSITE; PS00320; A4_INTRA; 1.
 DR PROSITE; PS00280; BPTI_KUNITZ_1; 1.
 DR PROSITE; PS00279; BPTI_KUNITZ_2; 1.
 DR Transmembrane; Alternative splicing; Serine protease inhibitor;
 KW Signal; Glycoprotein.
 FT SIGNAL 1 29
 FT CHAIN 30 765
 FT DOMAIN 30 695
 FT TRANSMEM 696 718
 FT DOMAIN 719 765
 FT DOMAIN 218 282
 FT DOMAIN 308 366
 FT ACT_SITE 322 323
 FT DISULFID 312 362
 FT DISULFID 321 345
 FT DISULFID 337 358
 FT DOMAIN 218 229
 FT CARBOHYD 628 628
 FT VARSPPLIC 311 365
 FT VARSPPLIC 616 627
 FT CONFLICT 575 577
 SQ SEQUENCE 765 AA; 86882 MW; CF51FCCCE305A0CF CRC64;
 Query Match 48.6%; Score 1971.5; DB 1; Length 765;
 Best Local Similarity 49.7%; Pred. No. 1.2e-95;
 Matches 401; Conservative 134; Mismatches 173; Indels 99; Gaps 20;
 QY 5 LALLLLAANTARALEV-----PTDGNAG---LLAEPQIAMFCGRLNMNMVQNGKWDSP 56
 DB 15 LVLLLLGLTAPAAALAGYIEALANAGTGFVAEPQIAMFCGKLMNMVNIQTGWEPDP 74
 QY 57 SGTCTIDTKEGILQCOEYYPELOITNVVEANQPVYTIQWCKRKRQCKTHPHFVYPR 116
 DB 75 TGTKSLGTKEVLQYCOEYYPELOITNVVEANQPVYTIQWCKRKRQCKTHPHFVYPR 132
 QY 117 CLVGFVSDALLVPDKCKFLHQRWDVCETHLHWHTVAKETCEKSTNLUHDYGMLLPGCI 176
 DB 133 CLVGFVSDVLLVPENCQFQHERMEVCCKHQHWHTVVKAECLTEGMYLYSGMLLPCGV 192
 QY 177 DKFRGVEFVCCPLAE--ESDNVDSADAEEDSDVWVGADTDYA-DGSEDKVVEVAEREE 233
 DB 193 DQFHGTGYVCCPTKVVDSDTMSKSEEEEB-----DEEDYALDKSEFTEADELDT 248
 QY 234 VAEEVEEADDDDDDDGDEVEEAEPEEYEE-----ATERITSATTTTTTESVEEYV 287
 DB 249 EAAADEDEDEEEEEEVEEDRYDYDFKGDYNEENPTSPSDGCTISDKETAHDV 308
 QY 288 REVCSQAEQTPGRAMISRWFDVTGKCAPFYGGCGGNRNNDTEECYMAVCSAMSQ 347
 DB 309 KAVCSQAEQTPGRAMVPRWYFVDSLKGKGVREFTYGGCGGNRNNDTEECYMAVCSMTI-- 366
 QY 348 SLKTTQEPPLARDPVKLPITTAASTPDVADVKYLETGPDENEHAFKAKERLEAKHERMS 407
 DB 367 -----PPTPEPT-----NDVDYFETSADNEHAFKAKERLEAKHERMS 408
 QY 408 QVMEWEAEERQAKNPKADKKAVIOHFQEKVESLEQEAANERQQLVETHMARVEAMLND 467
 DB 409 RVKKEWEAEERQAKNPKAERQTLQHFQAMVKALEKAASEKQQLVETHLARVEAMLND 468
 QY 468 RRLALENITALQAVPPRRHVFNMMLKKYVRAEQKDRQHTLKHFEHVRWMDPKAAQIR 527

DB 469 RRLALENITALQSDPPRPHRIQALRRYVRAENKORLHTIRHYQHVLAVDPEKAAQMK 528
 QY 528 SQVWTHLVIVYERNMQSLSLLYNYPVAAEETQDEVDLLOKEQNYSDVLANWLSSEPRIS 587
 DB 529 SQVWTHLVIVYERNMQSLSLLYNYPVAAEETQDEVDLLOKEQNYSDVLANWLSSEPRIS 574
 QY 588 YGNDALMPSLTETKTIVELLVNGEFSUDDLOQPHSFGADSVDPANTENEVEPVDARPAAD 647
 DB 575 ---DQFTSSISENPVDVR---VSSEES-EETPPPHFP--HFPFSLSENE----DTQPELY 621
 QY 648 RGLTTRPGSGLTN-----IKTEE---ISEVKMDAEFRHDSGYEVHHQKLVEFAEDVGS-- 697
 DB 622 HPM--KKGSGMAEQDGGGLGAEKVKNSKMDENMVIDELDV--KEMIFNAERVGGL 677
 QY 698 -----NKAIIGLMVGGVVIATVITVLMLKKKQYTSIHIGVYVEVDAV 742
 DB 678 EEPDSVGPLREDFSLSSSALIGLLVIAVIAIVISLVMLRKRQYGTISHIGVYVHPML 737
 QY 743 TPEERHLSKMQNGYENPTYKFFEQMQ 769
 DB 738 TPEERHLNKMONGYENPTYKYLEQMQ 764
 RESULT 7
 ID APP2_MOUSE STANDARD; PRT; 695 AA.
 AC O06335;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Amyloid-like protein 2 precursor (CDEI-box binding protein) (CDEBP).
 GN APLP2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fetal brain;
 RA von der Kammer H.;
 RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 1-246 FROM N.A.
 RX MEDLINE=94032480; PubMed=8218408;
 RA Hanes J., von der Kammer H., Kristjansson G.I., Scheit K.H.;
 RT "The complete cDNA coding sequence for the mouse CDEI binding
 protein.";
 RL Biochim. Biophys. Acta 1216:154-156(1993).
 RN [3]
 RP SEQUENCE OF 185-695 FROM N.A.
 RC STRAIN=BALB/c; TISSUE=Heart;
 RX MEDLINE=93129193; PubMed=1482349;
 RA Vidal F., Blangy A., Rassoulzadegan M., Cuzin F.;
 RT "A murine sequence-specific DNA binding protein shows extensive local
 similarities to the amyloid precursor protein.";
 RL Biochem. Biophys. Res. Commun. 189:1336-1341(1992).
 RN [4]
 RP SEQUENCE OF 1-35 FROM N.A.
 RC STRAIN=129/Sv;
 RX MEDLINE=96029629; PubMed=7592716;
 RA von Koch C.S., Lahiri D.K., Mammen A.L., Copeland N.G.,
 RA Gilbert D.J., Jenkins N.A., Sisodia S.S.;
 RT "The mouse APLP2 gene. Chromosomal localization and promoter
 characterization.";
 RL J. Biol. Chem. 270:25475-25480(1995).
 CC -1- FUNCTION: BINDS TO THE DNA 5'-GTCAATG-3' (CDEI BOX) WHICH PLAYS
 AN IMPORTANT ROLE IN THE EARLY DEVELOPMENT OF EMBRYOS.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN AND NUCLEAR
 (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE APP FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; 222592; CAA80306.1; -
 DR EMBL; M97216; AAA20039.1; -
 DR EMBL; U34291; AAC52318.1; -
 DR PIR; JC1404; JC1404.
 DR HSSP; P05067; 1MWP.
 DR MGD; MGI:88047; APLP2.
 DR InterPro; IPR001868; A4_APP.
 DR Pfam; PF02177; A4_EXTRA; 1.
 DR PRINTS; PR00203; AMYLOIDA4.
 DR SMART; SM00006; A4_EXTRA; 1.
 DR PROSITE; PS00319; A4_EXTRA; 1.
 DR PROSITE; PS00320; A4_INTRA; 1.
 KW Transmembrane; DNA-binding; Signal; Nuclear protein.
 FT SIGNAL 1 29
 FT CHAIN 30 695
 FT DOMAIN 30 624
 FT TRANSMEM 625 648
 FT TRANSMEM 649 695
 FT DOMAIN 218 294
 FT DOMAIN 218 231
 FT DOMAIN 256 266
 FT CARBOHYD 485 485
 FT CONFLICT 185 189
 SQ SEQUENCE 695 AA; 78944 MW; BBF4B95AAB2A0311 CRC64;

Query Match 42.0%; Score 1704.5; DB 1; Length 695;
 Best Local Similarity 46.9%; Pred. No. 8.5e-82;
 Matches 371; Conservative 113; Mismatches 170; Indels 137; Gaps 24;

QY 5 LALLLLAAMTARALEV-----PTDGNAG---LLAEPOIAMFCGRGLNMHNVQNGKWDSP 56
 DB 15 LVLLLLGLTAPAAALAGYTEALANAGTGFVAEPOIAMLCGLNMHNVQNGKWDSP 74
 QY 57 SGTKCTIDTKEGILQVCQEVYELQITNVVNEANQPTVQNWCKRCKQCKTHPHFVYPR 116
 DB 75 TGTKSCIGTKEEVLQVCQEVYELQITNVVNEANQPTVQNWCKRCKQCKTHPHFVYPR 132
 QY 117 CLNGFEVSDALLVPDKCKFLHQRMDVCEHLLHWHYVAKETCEKSTNLHDYGMLLPCGI 176
 DB 133 CLNGFEVSDVLLVPDNCQFFHQERMEVCEKHQRHWHLVKBEACLTGLTLYSGMLLPCGV 192
 QY 177 DKFRGEVFCVCLAEESDNVSDADAEEDSDVMWGGADTDYADGSDKVVVEVAEEVAE 236
 DB 193 DQFHGTEYVCCP---QTKTVDS-----DSTMSK-----EEEE--- 221
 QY 237 VEEDAEEDDEDGDEVEEAEPEEATERTSTATTITTTESVEEVREVCEQAE 296
 DB 222 -EEDAEEDDEDGDEVEEAEPEEATERTSTATTITTTESVEEVREVCEQAE 270
 QY 297 TGPCRAMISWYEDVTEGKCAPFFYCGCGNRNFDTEECYCAVCSAMSQSLLTKTQP 356
 DB 271 -----RDYYD-----PF-----KGDYNEEPT-----PSEGTI---SDKE 301
 QY 357 LARDPKLPTTAASTPAVDKYLETPGDENEHAFKAKERLEAKHRMSQVMREWEA 416
 DB 302 IVHD-VKVPPTPLTND-VDVILETSADNEHAFKAKERLEAKHRMSQVMREWEA 359
 QY 417 EROAKNLPKADKAVTOHQFQEKVESLEQEAANERQOLVETHMARVEAMLNDRRLALENY 476
 DB 360 ELQAKNLPKTERQTLQHQFQAVKALEKAASEKQQLVETHLARVEAMLNDRRLALENY 419
 QY 477 ITALQAVPPRPVFNKMLKYYVRAEQKDRQHTLKHFEHVWMDPKKAAQIRSQVMTLHV 536
 DB 420 LAALQSDPPRPHRILOALRRYVRAENKRLHTIRHQVLAVDPEKAAQKMSQVMTLHV 479
 QY 537 IYERMNQSLLLYNVPAVAEEIQDEVDLQKQNYSDVLANMISEPRISYGNDAIMP 596

DB 480 IEERNQSUULLYKVPYVAQEIQEIDELLQEQR-----ADM-----DQFTSS 522
 QY 597 LRETETTVELLVPNGEFSLDDLPWHSFGADSVPAANTENEVEPDARPAADRLTTRPGS 656
 DB 523 ISENPDVVRVSESE-EIPPPHPLHPF-----PSUSENE-----GSGMAEQDG- 565
 QY 657 GLTNKITEI-SEVKMDAEFRHDSGYEVHVKLVFAEDVGS-----N 698
 DB 566 GLIGAEKVKINSKMDENMVIDETLDV--KEMIFNAERVVGLEEPESVGPLREDEFSLS 623
 QY 699 KGAIIGLMVGGVVIATVITLVMLKKQYTSIHGVVEVDAAVTPEERHLSKMQQCYE 758
 DB 624 SNALIGLVIAVAIAIVIVISLVMLKKQYTSIHGVVEVDAAVTPEERHLSKMQQCYE 683
 QY 759 NPTYKFEQMQ 769
 DB 684 NPTYKLEQMQ 694
 RESULT 8
 ID APPL_HUMAN STANDARD; PRT; 650 AA.
 AC P51693; O00113;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Amyloid-like protein 1 precursor (APLP).
 GN APLP1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9808960; PubMed=9428684;
 RA Paliga K., Peraus G., Kregar S., Duwrrwang U., Hesse L., Muthaup G.,
 RA Masters C.L., Beyreuther K., Weidemann A.;
 RT "Human amyloid precursor-like protein 1--cDNA cloning, ectopic
 RT expression in COS-7 cells and identification of soluble forms in the
 RT cerebrospinal fluid.";
 RL Eur. J. Biochem. 250:354-363(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9818087; PubMed=9521588;
 RA Lenkkeri U., Kestila M., Lamerdin J., McCreedy P., Adamson A.,
 RA Olsen A., Tryggvason K.;
 RT "Structure of the human amyloid-precursor-like protein gene APLP1 at
 RT 19q13.1.";
 RL Hum. Genet. 102:192-196(1998).
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. C-TERMINALLY
 CC PROCESSED IN THE GOLGI COMPLEX AND IS THEN SECRETED.
 CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN BRAIN, PARTICULARLY
 CC IN THE CEREBRAL CORTEX POSTSYNAPTIC DENSITY.
 CC -!- PTM: N- AND O-GLYCOSYLATED.
 CC -!- SIMILARITY: BELONGS TO THE APP FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U48437; AAB96331.1; -
 DR EMBL; AD000864; AAB50173.1; -
 DR HSSP; P05067; 1MWP.
 DR Genew; HGNC:597; APLP1.
 DR MIM; 104775; -
 DR InterPro; IPR001868; A4_APP.
 DR Pfam; PF02177; A4_EXTRA; 1.
 DR PRINTS; PR00203; AMYLOIDA4.

```
DR SMART; SM00006; A4_EXTRA; 1.
DR PROSITE; PS00319; A4_EXTRA; 1.
DR PROSITE; PS00320; A4_INTRA; 1.
KW Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 21
FT CHAIN 22 650
FT DOMAIN 22 580
FT TRANSMEM 581 603
FT DOMAIN 604 650
FT DOMAIN 640 643
FT DOMAIN 241 247
FT DOMAIN 264 268
FT CARBOHYD 337 337
FT CARBOHYD 461 461
FT CARBOHYD 551 551
FT CONFLICT 48 48
FT SEQUENCE 650 AA; 72202 MW; 121A034B708C67CA CRC64;

Query Match 28.5%; Score 1150.5; DB 1; Length 650;
Best Local Similarity 35.7%; Pred. No. 2.9e-53;
Matches 279; Conservative 114; Mismatches 218; Indels 171; Gaps 20;

QY 1 MLPLGALLLL-----LAAWTARALEVPTDGNAGLLAEPQIAMFCGRLLNMHNVQNGKW 52
DB 23 LLPLLLLLRAQPAIGSLAGSGAPEAP--GSA-----QVAGLCGRLLTHRLDRTGRW 74
QY 53 DSDPSGFKTCDTKGELQVCEVPELQITNVVEANQPVTIQNKCKRCKQCKTHPHF- 111
DB 75 EPDQSRCLRPQVLEVCROMYPELQIARVEQATQALPWERWCGSGRSGSCAHPHQ 134
QY 112 VIPYRCILGVFEVDALVPDKKFLHQRMDVYCETHLHWHVTAKETCSKSTNLHDYGL 171
DB 135 WVPRLCPGFVSEALLVPCGRFLHQRMDQCESSTRRHOEAQACSSOGLLLHSGML 194
QY 172 LPCGIDFRGVFVCCPLAESDNVSADAEEDSDVWVGADTDYADGSEKVVVEAE 231
DB 195 LPGSDFRGVFVCCPPPTPD--PSGTAVGDPSTRW-----PPGSR---VEGAE 242
QY 232 EEVAEVEEEDDEDEDEDEVEEAEPEYEATERTTTATTTTTSVEEVREVC 291
DB 243 EE-----EESFPQVDDYFVEPPQAESEE----- 267
QY 292 SEQAETGCRAMISRYVDFVTEGKCAFFYGGCGGNRNFDEEYCMVCGSAMSLLK 351
DB 268 -----ETVP-----PPSSHTLAVVG-----K 283
QY 352 TTQEPLEARDVPLTTAATPDADVLYLETPGDENEHAHQAKERLEAKHRERMSQVR 411
DB 284 VT-----PTPRT-----DGVDFYGMGPGEISEHGFELRAKMDLEERRMQINEVMR 330
QY 412 EWEEARQAKNLKADKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRL 471
DB 331 EWAMADNQSNLKPADRAQLNEHFQSTLOTLEQVSGERQLVETHATRVIALINDORRA 390
QY 472 ALENYITALQVPPRPHVFNMLKVVYRAEQKDRQHTLKHFEHVRVMDPKKAAQIRSQVM 531
DB 391 ALGEFLAALQADPQABRVLLALRRYLRAEQKQRHTLRHYVAAVADPEKAQOMRFQVH 450
QY 532 THLRVYVERNQSLILLYNPVAAEETODEVDELLOKEQYNSDDVLNMISEPRISYND 591
DB 451 THLQVIEERNQSLGLLDQPHLAQELRPQIQLLSEH----- 489
QY 592 ALMPSLTETKTVELLPVNGEFLSD--DLQPHWSFGADSVPAANTEVEPVDARPAADR 649
DB 490 -LGPSELEA-----PAPGGSSEDKGLQPPDS--KDOTPM-----TLPKG 526
QY 650 LTRPGSGLNITKEETISEYKMDAETRHDSGYEVHH---QKLVFFAEDVGSNGKAIIGLM 706
DB 527 STEQDAASPEKEMKNPLEQYERKVNASVPRGFPFHSSEIQRDDELAPAGTGVSRVAVSGLL 586
QY 707 VGGVIVATVITLML-KKKQVTSHHGVVEVDAVTPPEERHLSKMOQNGYENPTYKFF 765
DB 587 IMGAGGSLIVLSMLLRKKPKYGAISHGVVEVDPMLTLEEQQQLRELQHRGYENPTYREL 646
```

```
QY 766 EQ 767
DB 647 EE 648

RESULT 9
APPL_MOUSE
ID APPL_MOUSE STANDARD; PRT; 653 AA.
AC Q03157;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE Amyloid-like protein 1 precursor (APLP).
GN APLP1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=93066322; PubMed=1279693;
RA Wasco W., Bupp K., Magendantz M., Gusella J.F., Tanzi R.E.,
RA Solomon F.;
RT "Identification of a mouse brain cDNA that encodes a protein related
RT to the Alzheimer disease-associated amyloid beta protein precursor.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:10758-10762(1992).
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. MAY BE PROCESSED
CC IN THE GOLGI COMPLEX; HOWEVER, IT IS NOT YET CLEAR WHETHER APLP
CC IS SECRETED.
CC -!- SIMILARITY: BELONGS TO THE APP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DB EMBL; L04538; AAA37247.1; -
DR PIR; A46362; A46362.
DR HSSP; P05067; LMWP.
DR MGD; MGI:88046; Aplp1.
DR InterPro; IPR001868; A4 App.
DR Pfam; PF02177; A4_EXTRA; 1.
DR PRINTS; PR00203; AMYLOIDA4.
DR SMART; SM00006; A4_EXTRA; 1.
DR PROSITE; PS00319; A4_EXTRA; 1.
DR PROSITE; PS00320; A4_INTRA; 1.
KW Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 20
FT CHAIN 21 653
FT DOMAIN 21 583
FT TRANSMEM 584 606
FT DOMAIN 607 653
FT DOMAIN 643 646
FT DOMAIN 263 271
FT CARBOHYD 464 464
FT CARBOHYD 554 554
FT SEQUENCE 653 AA; 72751 MW; 56516DC3EA40E4B0 CRC64;

Query Match 28.4%; Score 1150.5; DB 1; Length 653;
Best Local Similarity 35.4%; Pred. No. 5.3e-53;
Matches 274; Conservative 120; Mismatches 228; Indels 153; Gaps 19;

QY 1 MLPLGALLLLAAWTARA-LEVPTDGNAGLLAEPQIAMFCGRLLNMHNVQNGKWDSPSGT 59
DB 22 LLP-LSLLLLRAQLAVGNLVGSPSAEAPGSAQVAGLCGRLLTHRLDRTGRWEPDPQS 80
QY 60 KTCIDTKGILQVCEVPELQITNVVEANQPVTIQNKCKRCKQCKTHPHF-VIPYRL 118
```

```
Db 81 RCLLDQFVLECYROMYPELHIAVQAAQAIPIRMWCGTRSGRCAHPHEVVPFCHL 140
Qy 119 VGEVSDALLVPDKCKFLHQRMDVCEHLHWHHTVAKETSEKSTNLNHDYGMLLPGIDK 178
Db 141 PGEVSEALLVPECRLHQRMDVCEHLHWHHTVAKETSEKSTNLNHDYGMLLPGIDK 200
Qy 179 FRGVEVCCPLAEESDNVDSADAEEDSDVW-WGGADTDYADGSEDKVVEVAEEVAEV 237
Db 201 FRGVEVCCP-PPATPNPMAAGDPSTRSWPLGGR---AEGED-----EEVESF 248
Qy 238 EEEADDEDEDEDEVEBAEPEATERTTSIATTTTTTTESEVVEVVCSEQAET 297
Db 249 PQVDDYFVEPPQAESEEEERAPPSSHHP----- 281
Qy 298 GPCRAMISRVEDYTECKCAPFFYGGGGRNNFDTEECYMAVCGSAMQSLLKTTQEP 357
Db 282 ----VMSR-----VT----- 288
Qy 358 ARDVKLPTTAASPPDAVDKYLETGPDENEHAHFQKAKERLEAKHRMSQVMREWEAE 417
Db 289 ---PTPRPT-----DGVDFYFGMPGEIGEGHFLRAKMDLEERRMRQINEVNRWAMAD 339
Qy 418 ROAKNLKADKAVIQHFQKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYI 477
Db 340 SQSKNLKADRALNEHFQILQTEEQVSGERQRLVETHATRVIALINDQRAALEGFL 399
Qy 478 TALQAVPPRRHVFNMMLKKYVRAEQDROHTLKHFEHVRMVDPKKAAQIRSOVWTHLRY 537
Db 400 AALQGDPPQAEVLMALRYLRAEQEQRHTLRIYHVAADPEKACQMRQFQVOTHLQVI 459
Qy 538 YERNQSLSLYNVPAVEEIQDEVELLOKEQNYSDVLANNIMSEPRISYGNDAIMP-S 596
Db 460 EERNQSLGLDQNPQLAQLRQIQELL-----LAEHLGPSEL-----DASVPGS 505
Qy 597 LTKTKTVLLPNGEESLDLQPHWISFGADSVPAANTENEVEPVDARPAADRLTTPGS 656
Db 506 SSEDK-----GSLQP-----PEKDDPPVTLPL---KGSTQDESS 536
Qy 657 GLTNIKTEEISEVKMDAEFRHDSGYEVHH---OKLVFAEDVGSNKGAIIGLMVGWVIA 713
Db 537 SSGREKLTPLQEQVQKVNASAPGFFPHSHSDIORDELAPSGTGVSEALSGLLIMAGGG 596
Qy 714 TVIVITVLM-LKKQYTSIHGVVEVDAVAVTEERHLSKMQONGYENPTYKFEQ 767
Db 597 SLIVSLLLLRKKPKYGTISHGVVEVDPMLTLEEQRLQELQHGNYENPTYRFELE 651

RESULT 10
A4_CAEEL
ID A4_CAEEL STANDARD; PRT; 686 AA.
AC Q10651; Q18583; Q952X1;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Beta-amyloid-like protein precursor.
GN APL-1 OR C42D8.8.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OX Rhabditidae; Peloderinae; Caenorhabditis.
OC NCBI_TaxID=6239;
RN [1]
RP SEQUENCE OF 6-686 FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=94089766; PubMed=8265668;
RA Daigle I., Li C.;
RT "apl-1, a Caenorhabditis elegans gene encoding a protein related to
RL the human beta-amyloid protein precursor."
RL Proc. Natl. Acad. Sci. U.S.A. 90:12045-12049(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Hallsworth K.;
RL Submitted (May-1996) to the EMBL/GenBank/DBJ databases.
```

```
[3]
RN REVISIONS, AND ALTERNATIVE SPLICING.
RP Waterston R.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -!- ALTERNATIVE PRODUCTS: 2 isoforms; a and b (shown here); may be
CC produced by alternative splicing.
CC -!- SIMILARITY: BELONGS TO THE APP FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U00240; AAC46470.1; ALT_INIT.
CC EMBL; U56966; AAA98722.1; -.
CC EMBL; U56966; AAK68242.1; -.
CC HSP; P05067; MW.
CC WormPep; C42D8.8a; CE04209.
CC WormPep; C42D8.8b; CE27845.
CC InterPro; IPR001868; A4_APP.
CC Pfam; PF02177; A4_EXTRA; 1.
CC SMART; SM00006; A4_EXTRA; 1.
CC PROSITE; PS00319; A4_EXTRA; 1.
CC Signal; Transmembrane; Amyloid; Neurogenesis; Glycoprotein;
CC Alternative splicing.
CC SIGNAL 1 21
FT CHAIN 22 686
FT DOMAIN 22 621
FT TRANSMEM 622 642
FT DOMAIN 643 686
FT DOMAIN 205 228
FT DOMAIN 676 679
FT CARBOHYD 84 84
FT CARBOHYD 201 201
FT CARBOHYD 249 249
FT CARBOHYD 417 417
FT VARSPIC 538 539
SQ SEQUENCE 686 AA; 79434 MW; A0816858FDD48608 CRC64;
Query Match 19.3%; Score 785; DB 1; Length 686;
Best Local Similarity 26.8%; Pred. No. 6e-34;
Matches 224; Conservative 111; Mismatches 272; Indels 230; Gaps 23;
Qy 1 MLPGLALLLAATARALEVPTDGNAGLIAEPQIAMFCGRNMNMNVONGKWDSPSGTK 60
Db 6 LMIGLLIPLVA-TVVAEGSPAGSKRHEKFIPIWAFSCGYRQYM-TEEGSWKTDERYA 63
Qy 61 TCIDTKEGILQCYQEVPELQITNVVEANQPVTONWCKRGRKQCKTHPIHPIVRLVG 120
Db 64 TCFSGKLDILYKRAYPSMNITNIVEYSHEVSI SDWCREEGSPCK-WTHSVRPYHCIDG 122
Qy 121 ERVSDALLVPDKCKFLHQRMDVCEHLHWHHTVAKETSEKSTN-----LHDYGMLLPC 174
Db 123 EPHSEALQVPHDQCPSHNSRDCNDYQHWKDEAGCKTKSKGNKMDMVRSPFAVLPC 182
Qy 175 GIDKFRGVEVCCPLAEESDNVDSADAEEDSDVWGGADTDYADGSEDKVVEVAEEVE 234
Db 183 ALDMFTGVEFVCCP-----NDQTNKTDVQTK----- 209
Qy 235 AEVEEEADDDDEDEGEVEEEAEPEEATERTTSIATTTTTTTESEVVEVVCSEQ 294
Db 210 ---EDEDDEDDDAYEDDYSESEKDEE----- 236
Qy 295 AETGPCRAMISRWYFDVTEGKCAPFFYGGGGRNNFDTEECYMAVCGSAMQSLLKTTQ 354
Db 237 ----- 236
Qy 355 EPLARDPVKLPPTTAASPPDAVDKYLETGPDENEHAHFQKAKERLEAKHRMSQVMREWE 414
Db 355 EPLARDPVKLPPTTAASPPDAVDKYLETGPDENEHAHFQKAKERLEAKHRMSQVMREWE 414
```

Db 237 EPSQDP-----YFKIANWNEHDDFFKKAEMRMDEKHKRKKYVDKVMKENG 280

Qy 415 EA-----ERQAKNPKADKAVIQ---HFQEKVESLEQEAANERQOLVTHMARVEAMLN 466

Db 281 DLETRNEQKAD-PKGAEFKSQMARFQKTVSSLEEBHKMRKEIEAVHEERVQAMLN 339

Qy 467 DRRRLALENYITAL--QAVPPRPRHVFNNMLKKYVRAEQKDROHTLKHFEHVRMVPKKA 524

Db 340 EKRDATHDYQALATHVKNPKHNSVLQSLKAVIRAEKDRMHTLNRYRHLKADSKA 399

Qy 525 QIRSQVWTHLRVIERMNSLSLLYNP-----AVA--EEQDQVDELLQKQNY 573

Db 400 AYKPTVIHLRYIDLRLNGTAMLRDFFDLEKYVRPIAVTYWKDYRDESPDISVE---- 455

Qy 574 DDVLANNISEPRISYGN--DALMPSLT---ETKTTVELLPVNGEFSLDLQPHWSEFG 627

Db 456 DSELTPIIHDEFSKNAKLDVKAPTTAKPVKETNAKVLPTAEADSEAEDEYDEDD 515

Qy 628 SVPANT---ENEVEPYDARP-----AADRLGTTTPRGSGLTNKT 663

Db 516 EQVKKTPDMKKVKVVDIKPKEIKVTIEBEKAPKLVETSVQTDDEDDSSSTSS 575

Qy 664 EE-----ISEVKDAE-----FRHDSGYEVHHQKLVFAEDVGSNKGALIGLWGG 709

Db 576 DEDEDKNIKELRVDIETPIIDEPASFYRHD-----KLIQSPVEVSASSVFQPVVLAS 627

Qy 710 VVIATVITVLMLKKQVTSIHGVEVDAVTPPEHLSKMQONGYENPTYKFFE 766

Db 528 AMETAICTIAFAITNARRRRAMRGFIEVD-VYTPERHVAGQVNGYENPTYSFED 683

RESULT 11

A4_DROME STANDARD; PRT; 886 AA.

AC P14599;

DT 01-APR-1990 (Rel. 14, Created)

DT 01-APR-1990 (Rel. 14, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Beta-amyloid-like protein precursor.

GN APPL OR VND.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;

OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;

OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=89184650; PubMed=2494667;

RA Rosen D.R., Martin-Morris L., Luo L., White K.;

RT "A Drosophila gene encoding a protein resembling the human

RT beta-amyloid protein precursor.";

RL Proc. Natl. Acad. Sci. U.S.A. 86:2478-2482(1989).

RL [2]

RP SEQUENCE OF 1-83 FROM N.A.

RX MEDLINE=91184006; PubMed=2127912;

RA Martin-Morris L.E., White K.;

RT "The Drosophila transcript encoded by the beta-amyloid protein

RT precursor-like gene is restricted to the nervous system.";

RL Development 110:185-195(1990).

CC -1- FUNCTION: Probably corresponds to the protein encoded by the

CC essential locus vnd, a gene required for embryonic nervous system

CC development.

CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

CC -1- TISSUE SPECIFICITY: APPL TRANSCRIPTS ARE FOUND IN THE CENTRAL AND

CC PERIPHERAL NERVOUS SYSTEMS. WITHIN THE NERVOUS SYSTEM TRANSCRIPTS

CC ARE NOT OBSERVED IN NEUROBLASTS, NEWLY GENERATED NEURONS AND AT

CC LEAST ONE CLASS OF PRESUMED GLIAL CELLS.

CC -1- DEVELOPMENTAL STAGE: APPL TRANSCRIPTS ARE FOUND IN POST-MITOTIC

CC NEURONS IN THE CENTRAL AND PERIPHERAL NERVOUS SYSTEMS IN ALL

CC DEVELOPMENTAL STAGES.

CC -1- SIMILARITY: BELONGS TO THE APP FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL: J04516; AAA28874.1; -.

DR EMBL: X55774; CA339294.1; -.

DR EMBL: X55775; CA339294.1; JOINED.

DR PIR: A32758; A32758.

DR HSP: P05067; IWP.

DR Flybase: FBgn0000108; Appl.

DR InterPro: IPR001868; A4_APP.

DR Pfam: PF02177; A4_EXTRA; 1.

DR SMART: SM00006; A4_EXTRA; 1.

DR PROSITE: PS00319; A4_EXTRA; 1.

DR PROSITE: PS00320; A4_INTRA; 1.

KW SIGNAL; Transmembrane; Amyloid; Neurogenesis.

FT SIGNAL 1 27

FT CHAIN 28 886

FT DOMAIN 28 810

FT TRANSFEM 811 833

FT DOMAIN 834 886

FT DOMAIN 876 879

FT CARBOHYD 150 150

FT CARBOHYD 161 161

FT CARBOHYD 236 236

FT CARBOHYD 239 239

FT CARBOHYD 573 573

SQ SEQUENCE 886 AA; 98261 MW; C5C1EA4194DB5A9D CRC64;

Query Match 18.2%; Score 737.5; DB 1; Length 886;

Best Local Similarity 24.7%; Pred. No. 2.4e-31;

Matches 235; Conservative 137; Mismatches 313; Indels 265; Gaps 30;

Qy 7 LLLAAWTAARALEVPTDGNAGLLA-----EPQIAMFC--GRLNHMYV-QNGKWDSPSG 58

Db 9 LLLRLSLVWLAI-----GTAQVQAASPRWEPQIAVLCEAGIQPVLSSEGRWVTLSK 63

Qy 59 T---KTCIDTREGILOQCOEYVPELOITNVVNEANQPTVIONWCKRG---RKQCKTHPEV 112

Db 64 KTTGPTCLURDKMLLDYCKKAYPNRDIINIVESHYKIGWCGHOGALNAAKCKGSHRW 123

Qy 113 IPYRCLVGEFVSALLVPDKCKFLHQERMDVCEHHLHWHHTVAKETSEKSTNLNHDYGMLL 172

Db 124 KPFRL-CPFQSDALLVPEGLFDHIIHNASRCWPFVRWNTGAAACQERGQMFTFALL 182

Qy 173 PCGIDKFRGVFVCCP-----LAESDNVD-----SA 199

Db 183 PCGISVFGVEFVCCPKHFKTDEIHVKTKDLPVMPAAQINSANDELMDNDEDSNDSYK 242

Qy 200 DAEEDSDVWNGGADTDYADGSEDKVVEVAEEV-----AEV 237

Db 243 DANEDDDL-----DEDDLMDGDEEDDMVADEATAAGSPNTSGSGNSGSLDDINAEY 296

Qy 238 EE-EEADDEDEDEDEVEEAEPYEA-TERTTSTATTTTTTSTESVEEVVREVCSEA 295

Db 297 DSGEGDNYEEDGAGSEAEVEASWDGSGAKVVLKSDSSSPSSAPVAPKAPVKS 356

Qy 296 ETGPGRAMISRWYFDVTEGKCAPFFYGGCGGNRNNTFEYCMVAGCSMSQSLLKTTQE 355

Db 357 ESVTSTPQLS-----ASAAAFVAANSNGST-----GAGAPPSTAQPTSD 396

Qy 356 PLARDPVKLPITTAASTPDVADVKYLETPCDENEHAHFOKAKERLEAKHERMSQVNRWEE 415

Db 397 P-----YTFDPHYEHOSYKVSQARLEESREKRVTRVKNWDS 435

Qy 416 AERQAKNLPKADKKA-----VIQHFQEKVESLEQEAANERQOLVTHMARVEAMLNDR 468

Db 436 LEEKYQDMRLADPRAAKSFQKORTARFOTVSQALEEGNAEKHQLAAMHQORVLAHINOR 495

Qy 469 RRLALENITALQAVPPRPRHVRNMLKKYVRAEQKDROHTLKHFEH-VRWVDP---KKA 524

Db 496 KREAMTCVQALTEQPPNAHVCKLQKLLRALHOKRAHALAHYRHLNSGGPGGLEAAA 555
 QY 525 QIRSQVTHLRVIERWNOSLSLLYNNPVAEEI-----QDEV----- 562
 Db 556 SERPTLRLRIDRAVNSQMTMLKRYPELSAKIAQLMNDYIILALRSKDDIPGSSGLGMS 615
 QY 563 -----DELQKQBNYSDDVLAN 579
 Db 616 EAEAGILDKYRVEIKVAERKLEAKQKQRAERKLEAKKVDMDLKS 675
 QY 580 MISE-----PRISYGNDAIM-----PSLTETKTTVELLPVNG 611
 Db 676 QVABQSQPTQSSQSQAAQQQOEKSLPGKELGPDAAALVTAANPNLETTKS----- 726
 QY 612 EFSLDDLPQWHSFCADSPVANTENEVEPVDARPAADRGLTRPGSGLTNKTBEISEVKM 671
 Db 727 EKLSDFE-----YGEATVSTTKVQTVLPTVDDDAVRAVEDVAAA-----VAHQEA 773
 QY 672 DAEPFRHDSGYEVHHQKLVF-----FAEDVGSNK-----GAIIGLMVGGVVIATVIVITLVL 723
 Db 774 EPQVQHFWTHDLGHRSSFSLSLRREFAQHAAHAKGRNVYETLSFAGIALMAAEVGVAVA 833
 QY 724 KKKQYTSIH-HGVVEVDAAVTP-----EERHLSKMOQNGYENPTYKFFE 766
 Db 834 KWRTSRSPHAQGFTEVDQNVTHHPVIREKIVPNMQINGENPTYKYFE 883

RESULT 12

A4_MACFA STANDARD; PRT; 87 AA.
 ID A4_MACFA
 AC P53601;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Alzheimer's disease amyloid A4 protein (Fragment).
 GN APP.
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Macaca.
 OX NCBI_TaxID=9541;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91273117; PubMed=1905108;
 RA Podlinsky M.B., Tolan D.R., Selkoe D.J.;
 RT "Homology of the amyloid beta protein precursor in monkey and human
 RT supports a primate model for beta amyloidosis in Alzheimer's
 RT disease";
 RL Am. J. Pathol. 138:1423-1435(1991).
 CC -!- FUNCTION: FUNCTIONAL NEURONAL RECEPTOR WHICH COUPLES TO
 CC INTRACELLULAR SIGNALING PATHWAY THROUGH THE GTP-BINDING PROTEIN
 CC G(O) (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- ALTERNATIVE PRODUCTS: 5 ISOFORMS; APP(395), APP(563), APP(695),
 CC APP(751) AND APP(770) (SHOWN HERE); ARE PRODUCED BY ALTERNATIVE
 CC SPLICING (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE APP FAMILY.
 CC -!- SIMILARITY: CONTAINS 1 BPTI/KUNITZ INHIBITOR DOMAIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M58726; AAA36828.1; -
 CC HSP; P05067; IAAAP
 CC InterPro; IPR001868; A4_APP.
 CC InterPro; IPR002223; Kunitz_BPTI.
 CC Pfam; PF00014; Kunitz_BPTI; 1.

DR PRINTS: PR00759; BASICPTASE.
 DR ProDom: PD000222; Kunitz_BPTI; 1.
 DR SMART: SM00131; KO; 1.
 DR PROSITE: PS00319; A4_EXTRA; PARTIAL.
 DR PROSITE: PS00320; A4_INTRA; PARTIAL.
 DR PROSITE: PS00280; BPTI_KUNITZ_1; 1.
 DR PROSITE: PS00279; BPTI_KUNITZ_2; 1.
 KW Glycoprotein; Amyloid; Neurone; Alternative splicing;
 KW Serine protease inhibitor.
 FT NON_TER 1
 FT DOMAIN 4 79 BPTI/KUNITZ INHIBITOR.
 FT ACT_SITE 16 17 REACTIVE BOND.
 FT DISULFID 6 56 BY SIMILARITY.
 FT DISULFID 15 39 BY SIMILARITY.
 FT DISULFID 21 52 BY SIMILARITY.
 FT NON_TER 87 87
 SQ SEQUENCE 87 AA; 9608 MW; B67CG90DEE0EE7FF CRC64;
 Query Match 11.3%; Score 457; DB 1; Length 87;
 Best Local Similarity 95.4%; Pred No. 5, 6e-18;
 Matches 83; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 286 VREVCEQAETGPCRAMISRWFVDVTEGKCAPFFYGGCGGNRNNFDTEYCHAVCGSAM 345
 Db 1 VREVCEQAETGPCRAMISRWFVDVTEGKCAPFFYGGCGGNRNNFDTEYCHAVCGSVN 60
 QY 346 SQSLKTTQEPPLARDPVKLPPTAASTP 372
 Db 61 SQSLKTTREPLTRDPVKLPPTAASTP 87
 RESULT 13
 A4_MACMU
 ID A4_MACMU STANDARD; PRT; 76 AA.
 AC P29216;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Alzheimer's disease amyloid A4 protein (fragment).
 GN APP.
 OS Macaca mulatta (Rhesus macaque).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Macaca.
 OX NCBI_TaxID=9544;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RL Koo E.H., Sisodia S.S., Price D.L.;
 RL Submitted (JUL-1989) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: FUNCTIONAL NEURONAL RECEPTOR WHICH COUPLES TO
 CC INTRACELLULAR SIGNALING PATHWAY THROUGH THE GTP-BINDING PROTEIN
 CC G(O) (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- ALTERNATIVE PRODUCTS: 5 ISOFORMS; APP(395), APP(563), APP(695),
 CC APP(751) AND APP(770) (SHOWN HERE); ARE PRODUCED BY ALTERNATIVE
 CC SPLICING (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE APP FAMILY.
 CC -!- SIMILARITY: CONTAINS 1 BPTI/KUNITZ INHIBITOR DOMAIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X15985; CAA34116.1; -
 CC FIR; S06678; S06678.
 CC HSP; P05067; IAAAP.
 CC InterPro; IPR001868; A4_APP.
 CC InterPro; IPR002223; Kunitz_BPTI.

```

DR Pfam: PF00014; Kunitz_BPTI; 1.
DR ProDom: PD000222; Kunitz_BPTI; 1.
DR SMART: SM00131; KU; 1.
DR PROSITE: PS00319; A4_EXTRA; PARTIAL.
DR PROSITE: PS00320; A4_INTRA; PARTIAL.
DR PROSITE: PS00320; A4_INTRA; PARTIAL.
DR PROSITE: PS00280; BPTI_KUNITZ_1; 1.
DR PROSITE: PS00279; BPTI_KUNITZ_2; 1.
DR Glycoprotein; Amyloid; Neurone; Alternative splicing;
KW Serine protease inhibitor.
FT NON_TER 1 1
FT DOMAIN 1 76 BPTI/KUNITZ INHIBITOR.
FT ACT_SITE 13 14 REACTIVE BOND.
FT DISULFID 3 53 BY SIMILARITY.
FT DISULFID 12 36 BY SIMILARITY.
FT DISULFID 28 49 BY SIMILARITY.
FT NON_TER 76 76
FT SEQUENCE 76 AA; 8527 MW; 492BF3069AB082A1 CRC64;

Query Match 9.9%; Score 403; DB 1; Length 76;
Best Local Similarity 94.7%; Pred. No. 3.1e-15;
Matches 72; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 289 EVCSQETGCRAMISRWYFDVTEGKCAPFFYGGCGGNRNNFDTBEYCMVCGSAMSS 348
DB 1 EVCSQETGCRAMISRWYFDVTEGKCAPFFYGGCGGNRNNFDTBEYCMVCGSAMSS 60
QY 349 LLKTTQPLRPDKVL 364
DB 61 LRKTTREPLRPDKVL 76

RESULT 14
A4_BOVIN
ID A4_BOVIN STANDARD; PRT; 59 AA.
AC Q28033;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Alzheimer's disease amyloid A4 protein homolog [Contains: Beta-amyloid
protein (Beta-APP) (A-beta)] (Fragment).
GN APP.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=92017079; PubMed=1656157;
RA Johnstone E.M., Chaney M.O., Norris F.H., Pascual R., Little S.P.;
RT "Conservation of the sequence of the Alzheimer's disease amyloid
peptide in dog, polar bear and five other mammals by cross-species
polymerase chain reaction analysis.";
RL Brain Res. Mol. Brain Res. 10:299-305(1991).
CC -!- FUNCTION: FUNCTIONAL NEURONAL RECEPTOR WHICH COUPLES TO
INTRACELLULAR SIGNALING PATHWAY THROUGH THE GTP-BINDING PROTEIN
G(O) (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: BELONGS TO THE APP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X56124; CAA39589.1; -
CC DR EMBL; X56126; CAA39591.1; -
CC HSSP; P05067; IBA4.
CC InterPro: IPR001868; A4_APP.

DR Pfam: PF00014; Kunitz_BPTI; 1.
DR ProDom: PD000222; Kunitz_BPTI; 1.
DR SMART: SM00131; KU; 1.
DR PROSITE: PS00319; A4_EXTRA; PARTIAL.
DR PROSITE: PS00320; A4_INTRA; PARTIAL.
DR PROSITE: PS00320; A4_INTRA; PARTIAL.
DR PROSITE: PS00280; BPTI_KUNITZ_1; 1.
DR PROSITE: PS00279; BPTI_KUNITZ_2; 1.
DR Glycoprotein; Amyloid; Neurone; Alternative splicing;
KW Serine protease inhibitor.
FT NON_TER 1 1
FT DOMAIN 1 76 BPTI/KUNITZ INHIBITOR.
FT ACT_SITE 13 14 REACTIVE BOND.
FT DISULFID 3 53 BY SIMILARITY.
FT DISULFID 12 36 BY SIMILARITY.
FT DISULFID 28 49 BY SIMILARITY.
FT NON_TER 76 76
FT SEQUENCE 76 AA; 8527 MW; 492BF3069AB082A1 CRC64;

Query Match 9.9%; Score 403; DB 1; Length 76;
Best Local Similarity 94.7%; Pred. No. 3.1e-15;
Matches 72; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 289 EVCSQETGCRAMISRWYFDVTEGKCAPFFYGGCGGNRNNFDTBEYCMVCGSAMSS 348
DB 1 EVCSQETGCRAMISRWYFDVTEGKCAPFFYGGCGGNRNNFDTBEYCMVCGSAMSS 60
QY 349 LLKTTQPLRPDKVL 364
DB 61 LRKTTREPLRPDKVL 76

RESULT 14
A4_BOVIN
ID A4_BOVIN STANDARD; PRT; 59 AA.
AC Q28033;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Alzheimer's disease amyloid A4 protein homolog [Contains: Beta-amyloid
protein (Beta-APP) (A-beta)] (Fragment).
GN APP.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=92017079; PubMed=1656157;
RA Johnstone E.M., Chaney M.O., Norris F.H., Pascual R., Little S.P.;
RT "Conservation of the sequence of the Alzheimer's disease amyloid
peptide in dog, polar bear and five other mammals by cross-species
polymerase chain reaction analysis.";
RL Brain Res. Mol. Brain Res. 10:299-305(1991).
CC -!- FUNCTION: FUNCTIONAL NEURONAL RECEPTOR WHICH COUPLES TO
INTRACELLULAR SIGNALING PATHWAY THROUGH THE GTP-BINDING PROTEIN
G(O) (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: BELONGS TO THE APP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X56124; CAA39589.1; -
CC DR EMBL; X56126; CAA39591.1; -
CC HSSP; P05067; IBA4.
CC InterPro: IPR001868; A4_APP.

```

```

DR InterPro: IPR001255; Beta-APP.
DR Pfam: PF03494; Beta-APP; 1.
DR PROSITE: PS00319; A4_EXTRA; PARTIAL.
DR PROSITE: PS00320; A4_INTRA; PARTIAL.
DR Glycoprotein; Amyloid; Neurone; Transmembrane.
FT NON_TER 1 1
FT CHAIN 7 49 BETA-AMYLOID PROTEIN (POTENTIAL).
FT DOMAIN <1 34 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 35 58 POTENTIAL.
FT DOMAIN 59 >59 CYTOPLASMIC (POTENTIAL).
FT NON_TER 59 59
FT SEQUENCE 59 AA; 6414 MW; F43469D488A2E12D CRC64;

Query Match 7.2%; Score 292; DB 1; Length 59;
Best Local Similarity 100.0%; Pred. No. 1.3e-09;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

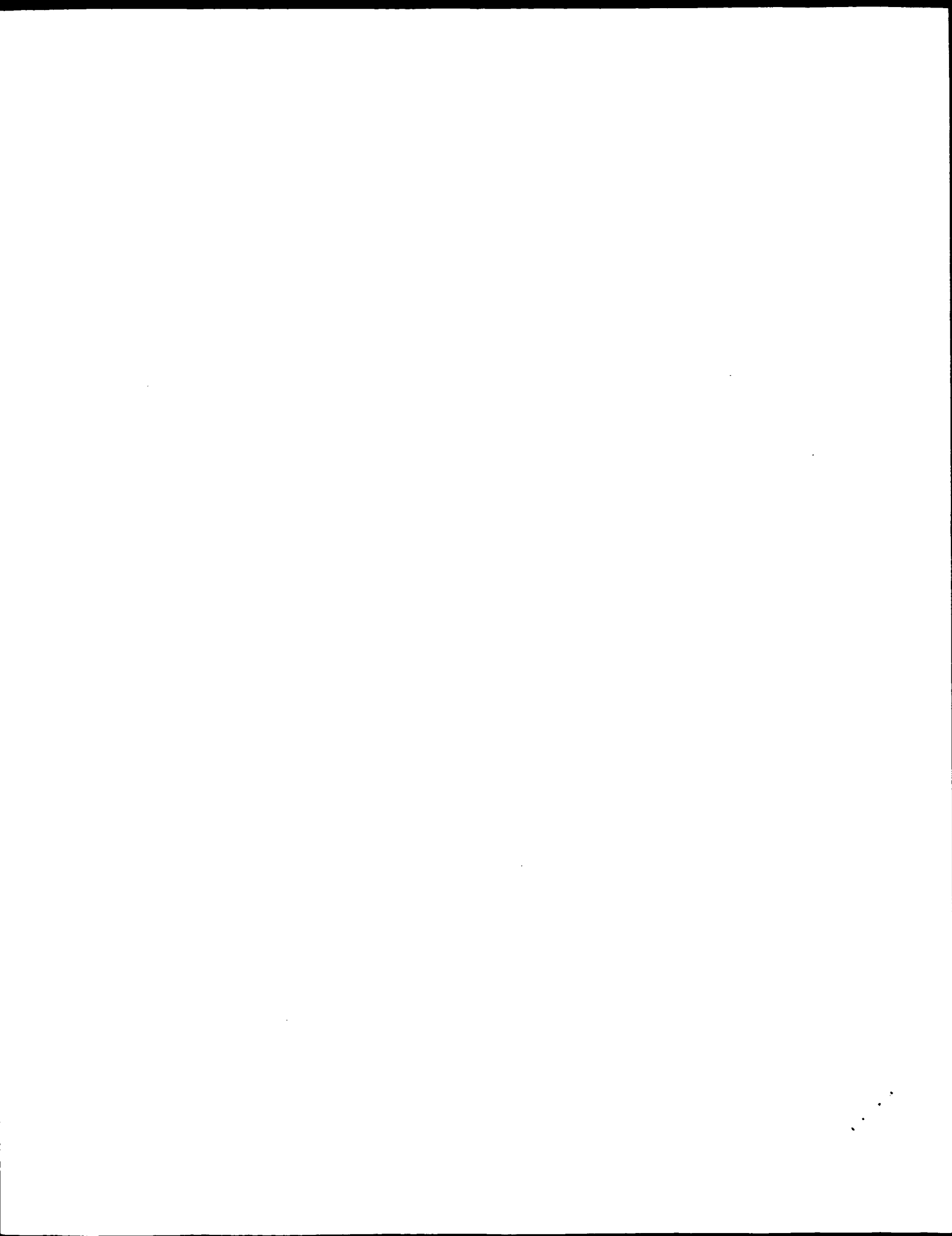
QY 666 ISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIAVITLVMLK 724
DB 1 ISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIAVITLVMLK 59

RESULT 15
A4_RABIT
ID A4_RABIT STANDARD; PRT; 58 AA.
AC Q28748;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Alzheimer's disease amyloid A4 protein homolog [Contains: Beta-amyloid
protein (Beta-APP) (A-beta)] (Fragment).
GN APP.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=92017079; PubMed=1656157;
RA Johnstone E.M., Chaney M.O., Norris F.H., Pascual R., Little S.P.;
RT "Conservation of the sequence of the Alzheimer's disease amyloid
peptide in dog, polar bear and five other mammals by cross-species
polymerase chain reaction analysis.";
RL Brain Res. Mol. Brain Res. 10:299-305(1991).
CC -!- FUNCTION: FUNCTIONAL NEURONAL RECEPTOR WHICH COUPLES TO
INTRACELLULAR SIGNALING PATHWAY THROUGH THE GTP-BINDING PROTEIN
G(O) (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: BELONGS TO THE APP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X56129; CAA39594.1; -
CC HSSP; P05067; IBA4.
CC InterPro: IPR001868; A4_APP.
CC InterPro: IPR001255; Beta-APP.
DR Pfam: PF03494; Beta-APP; 1.
DR PROSITE: PS00319; A4_EXTRA; PARTIAL.
DR PROSITE: PS00320; A4_INTRA; PARTIAL.
DR Glycoprotein; Amyloid; Neurone; Transmembrane.
FT NON_TER 1 1
FT CHAIN 6 48 BETA-AMYLOID PROTEIN (POTENTIAL).
FT DOMAIN <1 33 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 34 57 POTENTIAL.
FT DOMAIN 58 >58 CYTOPLASMIC (POTENTIAL).
FT NON_TER 58 58

```

SQ SEQUENCE 58 AA; 6300 MW; F434209D88EBA82D CRC64;
 Query Match 7.1%; Score 288; DB 1; Length 58;
 Best Local Similarity 100.0%; Pred. No. 2.1e-09;
 Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 667 SEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVIATVIVITLVMLK 724
 Db 1 SEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVIATVIVITLVMLK 58

Search completed: January 28, 2003, 14:14:59
 Job time : 15.7782 secs



GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:04:57 ; Search time 38.9101 Seconds
(without alignments)
4077.505 Million cell updates/sec

Title: US-09-904-987-2

Perfect score: 4058

Sequence: 1 MLPGLALLLAATAWTALEV.....KMQQNGYENPTYKFFEQMQN 770

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3981	98.1	770	6 Q9TUI0	Q9tui0 sus scrofa
2	3719.5	91.7	751	13 Q9DGJ7	Q9dgj7 gallus gall
3	3590.5	88.5	695	6 Q95KN7	Q95kn7 macaca fasc
4	3522.5	86.8	695	11 Q60496	Q60496 cavia sp. p
5	3489.5	86.0	695	11 P97487	P97487 mus musculus
6	3403.5	83.9	747	13 Q91963	Q91963 xenopus ap
7	3377.5	83.2	695	13 Q9DGJ8	Q9dgj8 gallus gall
8	3163.5	78.0	693	13 Q98SG0	Q98sg0 xenopus lae
9	3152.5	77.7	695	13 Q98SF9	Q98sf9 xenopus lae
10	3065	75.5	607	11 Q99K32	Q99k32 mus musculus
11	2914	71.8	699	13 Q57394	Q57394 narke japon
12	2831	69.8	547	4 Q13764	Q13764 homo sapien
13	2785.5	68.6	738	13 Q90W28	Q90w28 brachydanio
14	2763.5	68.1	780	13 Q73683	Q73683 tetraodon f
15	2717	67.0	569	13 Q9PVL1	Q9pvl1 gallus gall
16	2699	66.5	737	13 Q93279	Q93279 fugu rubrip

17	2562.5	63.1	534	13	Q93296	Q93296 gallus gall
18	2446.5	60.3	484	4	Q13793	Q13793 homo sapien
19	2437	60.1	694	13	Q8UUR9	Q8uur9 brachydanio
20	2288.5	56.4	612	13	Q919E7	Q919e7 brachydanio
21	1960.5	48.3	763	11	Q61482	Q61482 mus musculus
22	1956.5	48.2	751	11	Q60709	Q60709 mus musculus
23	1733.5	42.7	695	4	Q13861	Q13861 homo sapien
24	1721	42.4	669	4	Q14662	Q14662 homo sapien
25	1709.5	42.1	695	11	Q64348	Q64348 mus musculus
26	1604.5	39.5	472	13	Q8UUS0	Q8uus0 brachydanio
27	1569	38.7	305	4	Q9BT38	Q9bt38 homo sapien
28	1350.5	33.3	357	13	Q8UUI8	Q8uii8 brachydanio
29	1302	32.1	523	4	Q14594	Q14594 homo sapien
30	1267	31.2	522	4	Q9BT36	Q9bt36 homo sapien
31	1159.5	28.6	650	4	Q96A32	Q96a32 homo sapien
32	1150.5	28.4	654	11	Q8VC38	Q8vc38 mus musculus
33	788	19.4	160	11	Q9QZ78	Q9qz78 cavia sp. p
34	771	19.0	239	13	Q8UUI7	Q8uii7 brachydanio
35	739	18.2	887	5	Q9U4H3	Q9u4h3 drosophila
36	738	18.2	887	5	Q9TVV0	Q9tvv0 drosophila
37	678	16.7	136	6	P79307	P79307 sus scrofa
38	661	16.3	816	5	Q9W5F1	Q9w5f1 drosophila
39	561	13.8	182	11	Q9CYS4	Q9cys4 mus musculus
40	493	12.1	97	4	Q13778	Q13778 homo sapien
41	478	11.8	97	6	Q28673	Q28673 oryctolagus
42	411	10.1	82	4	P78438	P78438 homo sapien
43	402	9.9	75	4	Q9UQS8	Q9uqs8 homo sapien
44	393.5	9.7	82	4	Q16019	Q16019 homo sapien
45	389.5	9.6	82	4	Q16014	Q16014 homo sapien

ALIGNMENTS

RESULT 1

Q9TUI0 PRELIMINARY; PRT; 770 AA.
 AC Q9TUI0;
 DT 01-MAY-2000 (TREMREL. 13, Created)
 DT 01-MAY-2000 (TREMREL. 13, Last sequence update)
 DT 01-JUN-2002 (TREMREL. 21, Last annotation update)
 DE Amyloid precursor protein.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_taxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kimura A., Takahashi T.;
 RT "Amyloid Precursor Protein 770".
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB032550; BAA84580.1; -.
 DR HSSP; P05067; IAAP.
 DR InterPro; IPR001868; A4_APP.
 DR InterPro; IPR001255; Beta-APP.
 DR InterPro; IPR002223; Kunitz_BPTI.
 DR Pfam; PF02177; A4_EXTRA; 1.
 DR Pfam; PF03494; Beta-APP; 1.
 DR Pfam; PF00014; Kunitz_BPTI; 1.
 DR PRINTS; PR00203; AMYLOIDA4.
 DR PRINTS; PR00759; BASICPTASE.
 DR ProDom; PD000222; Kunitz_BPTI; 1.
 DR SMART; SM00006; A4_EXTRA; 1.
 DR SMART; SM00131; KU; 1.
 DR PROSITE; PS00319; A4_EXTRA; 1.
 DR PROSITE; PS00320; A4_INTRA; 1.
 DR PROSITE; PS00280; BPTI_KUNITZ_1; 1.
 DR PROSITE; PS50279; BPTI_KUNITZ_2; 1.
 KW Serine protease inhibitor.
 SQ SEQUENCE 770 AA; 86961 MW; 5F7A1DCB2BCC583E CRC64;
 Query Match 98.1%; Score 3981; DB 6; Length 770;
 Best Local Similarity 97.8%; Pred. No. 6.1e-232;

```

Matches 753; Conservative 8; Mismatches 9; Indels 0; Gaps 0;
Qy 1 MLPGLALLLAATAARALEVPTDGNAGLLAEPOIAMFCGRNLNMHNVQNGKWDSPSGTK 60
Db 1 MLPGLALLLAATAARALEVPTDGNAGLLAEPOIAMFCGRNLNMHNVQNGKWDSPSGTK 60
Qy 61 TCIDTKGILQYQCEVYPELQITNVVEANQPVITQNWCKRGKQCKTHPHFVPIYRCLVG 120
Db 61 TCIDTKGILQYQCEVYPELQITNVVEANQPVITQNWCKRGKQCKTHPHFVPIYRCLVG 120
Qy 121 EFVSDALLVPDKCFKFLHQERMDVCETHLHWHITVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
Db 121 EFVSDALLVPDKCFKFLHQERMDVCETHLHWHITVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
Qy 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEAEVEEAEVEE 240
Db 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEAEVEEAEVEE 240
Qy 241 EADDEDEDEGDEVEEAEPEEATERTTSIATTTTTTETESVEEVREVCSQAETGPC 300
Db 241 EADDEDEDEGDEVEEAEPEEATERTTSIATTTTTTETESVEEVREVCSQAETGPC 300
Qy 301 RAMISRWYFDVTGKCAPFYGGCGGNRRNFDTEECMAVCGSAMQSLLKTTQEPFLARD 360
Db 301 RAMISRWYFDVTGKCAPFYGGCGGNRRNFDTEECMAVCGSAMQSLLKTTQEPFLARD 360
Qy 361 PVKLPKADKAVIOHFQKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENIT 420
Db 361 PVKLPKADKAVIOHFQKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENIT 420
Qy 421 KNLPLKADKAVIOHFQKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENIT 480
Db 421 KNLPLKADKAVIOHFQKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENIT 480
Qy 481 QAVPPRRPRHVFNLKMYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 540
Db 481 QAVPPRRPRHVFNLKMYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 540
Qy 541 MNQSLSLYNNVPAVEEIQDEVDLQKEQNYSDVLANNISPRISYGNDAIMPSTET 600
Db 541 MNQSLSLYNNVPAVEEIQDEVDLQKEQNYSDVLANNISPRISYGNDAIMPSTET 600
Qy 601 KTTVELLPVNGEFLDLPQWHSFGADSVDPANTENEVEPVDARPAADRLGTLTPGSGLTN 660
Db 601 KTTVELLPVNGEFLDLPQWHSFGADSVDPANTENEVEPVDARPAADRLGTLTPGSGLTN 660
Qy 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFPAEDVGSNKGAIIGLVMGVVVIATVITL 720
Db 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFPAEDVGSNKGAIIGLVMGVVVIATVITL 720
Qy 721 VMLKKQYTSIHGCVVEVDAVTPPEERHLSKMQONGYENPTYKFFEQMQN 770
Db 721 VMLKKQYTSIHGCVVEVDAVTPPEERHLSKMQONGYENPTYKFFEQMQN 770

RESULT 2
Q9DGJ7 PRELIMINARY; PRT; 751 AA.
AC Q9DGJ7;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Beta-amyloid precursor protein 751 isoform.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Sarasa M., Rodolase A., Sorribas V.;
RT "Cloning of full-length chicken beta-amyloid precursor protein
isoforms."

```

```

RL Submitted (JUL-2000) to the EMBL/genBank/DBJ databases.
DR EMBL; AF289219; AAG00594.1; -.
DR HSSP; P05067; 1BA4.
DR InterPro; IPR001868; A4_APP.
DR InterPro; IPR001255; Beta-APP.
DR InterPro; IPR002223; Kunitz_BPTI.
DR Pfam; PF02177; A4_EXTRA; 1.
DR Pfam; PF03494; Beta-APP; 1.
DR Pfam; PF00014; Kunitz_BPTI; 1.
DR PRINTS; PR00203; AMYLOIDA4.
DR PRINTS; PR00759; BASICPTASE.
DR ProDom; PD000222; Kunitz_BPTI; 1.
DR SMART; SM00006; A4_EXTRA; 1.
DR SMART; SM00131; KU; 1.
DR PROSITE; PS00319; A4_EXTRA; 1.
DR PROSITE; PS00320; A4_INTRA; 1.
DR PROSITE; PS00280; BPTI_KUNITZ_1; 1.
DR PROSITE; PS00279; BPTI_KUNITZ_2; 1.
KW Serine protease inhibitor.
SQ SEQUENCE 751 AA; 84705 MW; E78E9413A8033D84 CRC64;

```

```

Query Match 91.7%; Score 3719.5; DB 13; Length 751;
Best Local Similarity 91.7%; Pred. No. 3.5e-216;
Matches 708; Conservative 18; Mismatches 23; Indels 23; Gaps 4;

```

```

Qy 1 MLPGLALLLAATAARALEVPTDGNAGLLAEPOIAMFCGRNLNMHNVQNGKWDSPSGTK 60
Db 1 MLPGLALLLAATAARALEVPTDGNAGLLAEPOIAMFCGRNLNMHNVQNGKWDSPSGTK 60
Qy 61 TCIDTKGILQYQCEVYPELQITNVVEANQPVITQNWCKRGKQCKTHPHFVPIYRCLVG 120
Db 61 TCIDTKGILQYQCEVYPELQITNVVEANQPVITQNWCKRGKQCKTHPHFVPIYRCLVG 120
Qy 121 EFVSDALLVPDKCFKFLHQERMDVCETHLHWHITVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
Db 121 EFVSDALLVPDKCFKFLHQERMDVCETHLHWHITVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
Qy 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEAEVEEAEVEE 238
Db 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEAEVEEAEVEE 240
Qy 239 EEEADDEDEDEGDEVEEAEPEEATERTTSIATTTTTTETESVEEVREVCSQAETG 298
Db 241 DEEADDD-DEDDGDI-BETEEYEAEATERTTSIATTTTTTETESVEEVREVCSQAETG 298
Qy 299 PCRAMISRWYFDVTGKCAPFYGGCGGNRRNFDTEECMAVCGSAMQSLLKTTQEPFLA 358
Db 299 PCRAMISRWYFDVTGKCAPFYGGCGGNRRNFDTEECMAVCGSV----- 344
Qy 359 RDPVKLPTTAASTPDVADVKYLETGPDENEHAFQKAKERLEAKHRERMSQVMREWEAE 418
Db 345 -----LPTTAASTPDVADVKYLETGPDENEHAFQKAKERLEAKHRERMSQVMREWEAE 399
Qy 419 QAKNLPKADKAVIOHFQKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENIT 478
Db 400 QAKNLPKADKAVIOHFQKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENIT 459
Qy 479 ALQAVPPRRPRHVFNLKMYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVI 538
Db 460 ALQAVPPRRPRHVFNLKMYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVI 519
Qy 539 ERMNQSLSLYNNVPAVEEIQDEVDLQKEQNYSDVLANNISPRISYGNDAIMPST 598
Db 520 ERMNQSLSFLYNNVPAVEEIQDEVDLQKEQNYSDVLANNISPRISYGNDAIMPST 579
Qy 599 ETKTTVELLPVNGEFLDLPQWHSFGADSVDPANTENEVEPVDARPAADRLGTLTPGSG 658
Db 580 ETKTTVELLPVNGEFLDLPQWHSFGADSVDPANTENEVEPVDARPAADRLGTLTPGSG 639
Qy 659 TNIKTEEISEVKMDAEFRHDSGYEVHHQKLVFPAEDVGSNKGAIIGLVMGVVVIATV 718
Db 640 TNVKTTEESEVKMDAEFRHDSGYEVHHQKLVFPAEDVGSNKGAIIGLVMGVVVIATV 699

```

QY 719 TLVLMKKKQYTSIHGGVVEVDAVTPERHLSKMQONGYENPTYKFEQMQN 770
 Db 700 TLVLMKKKQYTSIHGGVVEVDAVTPERHLSKMQONGYENPTYKFEQMQN 751

RESULT 3

Q95KN7
 ID Q95KN7 PRELIMINARY; PRT; 695 AA.
 AC Q95KN7
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-JUN-2002 (TrEMBLrel. 19, Last sequence update)
 DE Amyloid b-protein precursor.
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Macaca.
 OX NCBI_TaxID=9541;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=CEREBELLUM;
 RX MEDLINE=91273117; PubMed=1905108;
 RA Podlinsky M.B., Tolan D.R., Selkoe D.J.;
 RT "Homology of the amyloid beta protein precursor in monkey and human
 RT supports a primate model for beta amyloidosis in Alzheimer's
 RT disease.";
 RL Am. J. Pathol. 138:1423-1435(1991).
 DR EMBL; M58727; AAA36829.1; -;
 DR InterPro; IPR001868; A4_APP.
 DR InterPro; IPR001255; Beta-APP.
 DR Pfam; PF02177; A4_EXTRA; 1.
 DR Pfam; PF03494; Beta-APP; 1.
 DR PROSITE; PS00319; A4_EXTRA; UNKNOWN_1.
 DR PROSITE; PS00320; A4_INTRA; UNKNOWN_1.
 FT SIGNAL 1
 FT CHAIN 17 POTENTIAL.
 FT CHAIN 597 636 POTENTIAL.
 SQ SEQUENCE 695 AA; 78663 MW; 4F6EA0139F969D56 CRC64;

Query Match 88.5%; Score 3590.5; DB 6; Length 695;
 Best Local Similarity 90.1%; Pred. No. 1.9e-208;
 Matches 694; Conservative 1; Mismatches 0; Indels 75; Gaps 1;

QY 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPOIAMFCGRNLNHNHNQNGKWDSPSGTK 60
 Db 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPOIAMFCGRNLNHNHNQNGKWDSPSGTK 60
 QY 61 TCIDTKEGILQYCEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVYRCLVG 120
 Db 61 TCIDTKEGILQYCEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVYRCLVG 120
 QY 121 EFVSDALLVPDKCKFLQOERMDVCETHLHWHHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
 Db 121 EFVSDALLVPDKCKFLQOERMDVCETHLHWHHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
 QY 181 GVEFVCCPLAESDNVDSADAEDSDVWVGADTDYADGSDKVVEVAEEVAEVEE 240
 Db 181 GVEFVCCPLAESDNVDSADAEDSDVWVGADTDYADGSDKVVEVAEEVAEVEE 240
 QY 241 EADDDDDDDGDEVEEAEPEYEATERTSTATTATTTTTSVEEVREVCSQAETGPC 300
 Db 241 EADDDDDDDGDEVEEAEPEYEATERTSTATTATTTTTSVEEVREVCSQAETGPC 300
 QY 301 RAVISRWYFDVTGKCAPFFYGGCGGNRNFTTEYCHAVCGSAMSQSLLKTTQEPPLARD 360
 Db 289 ----- 288
 QY 361 PVKLPPTAASPDVADKYLETGDENEHAHFOKAKERLEAKHRERMSQVMREWEAEERQA 420
 Db 289 ---VPTAASPDVADKYLETGDENEHAHFOKAKERLEAKHRERMSQVMREWEAEERQA 345
 QY 421 KNLPRADKAVTQHFQEKVESLEQEAANERQOLVETHMARVEAMLNDRRLALENYITAL 480
 Db 346 KNLPRADKAVTQHFQEKVESLEQEAANERQOLVETHMARVEAMLNDRRLALENYITAL 405

QY 481 QAVPPRPHVENMLKKYVRABQKDRQHTLKHFEHVRVMDPKAAQIRSQVWTHLRVYER 540
 Db 406 QAVPPRPHVENMLKKYVRABQKDRQHTLKHFEHVRVMDPKAAQIRSQVWTHLRVYER 465
 QY 541 MNQSLSLLYNVPAAVEETQDEVDDELLOKEQNYSDVLANMISEPRIISYGNDAIMPSTET 600
 Db 466 MNQSLSLLYNVPAAVEETQDEVDDELLOKEQNYSDVLANMISEPRIISYGNDAIMPSTET 525
 QY 601 KTTVELLPVNGEFSLLDQLQPMHSEFADSPANTENEPEVDARPAADRGTLTTRPGSLTN 660
 Db 526 KTTVELLPVNGEFSLLDQLQPMHSEFADSPANTENEPEVDARPAADRGTLTTRPGSLTN 585
 QY 661 IKTEISEVKMDAEFRHDSGYEVHVKLVFAEDVGSNGKGAIIGLMVGGVVATVITL 720
 Db 586 IKTEISEVKMDAEFRHDSGYEVHVKLVFAEDVGSNGKGAIIGLMVGGVVATVITL 645
 QY 721 VMLKKKQYTSIHGGVVEVDAVTPERHLSKMQONGYENPTYKFEQMQN 770
 Db 646 VMLKKKQYTSIHGGVVEVDAVTPERHLSKMQONGYENPTYKFEQMQN 695

RESULT 4

Q60496
 ID Q60496 PRELIMINARY; PRT; 695 AA.
 AC Q60496
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Putative amyloid precursor protein.
 OS Cavia sp.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
 OX NCBI_TaxID=10143;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BRAIN;
 RX MEDLINE=97236426; PubMed=9116031;
 RA Beck M., Mueller D., Bigl V.;
 RT "Amyloid precursor protein in Guinea pigs - complete cDNA sequence and
 RT alternative splicing.";
 RL Biochim. Biophys. Acta 1351:17-21(1997).
 DR EMBL; X97631; CAA66230.1; -;
 DR HSP; P05067; IBA4
 DR InterPro; IPR001868; A4_APP.
 DR InterPro; IPR001255; Beta-APP.
 DR Pfam; PF02177; A4_EXTRA; 1.
 DR Pfam; PF03494; Beta-APP; 1.
 DR PRINTS; PR00203; AMYLOIDA4.
 DR SMART; SM00006; A4_EXTRA; 1.
 DR PROSITE; PS00319; A4_EXTRA; 1.
 DR PROSITE; PS00320; A4_INTRA; 1.
 SQ SEQUENCE 695 AA; 78701 MW; 5196A0C4017F16AB CRC64;

Query Match 86.8%; Score 3522.5; DB 11; Length 695;
 Best Local Similarity 88.2%; Pred. No. 2.4e-204;
 Matches 679; Conservative 7; Mismatches 9; Indels 75; Gaps 1;

QY 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPOIAMFCGRNLNHNHNQNGKWDSPSGTK 60
 Db 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPOIAMFCGRNLNHNHNQNGKWDSPSGTK 60
 QY 61 TCIDTKEGILQYCEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVYRCLVG 120
 Db 61 TCIGSGEILQYCEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVYRCLVG 120
 QY 121 EFVSDALLVPDKCKFLQOERMDVCETHLHWHHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
 Db 121 EFVSDALLVPDKCKFLQOERMDVCETHLHWHHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
 QY 181 GVEFVCCPLAESDNVDSADAEDSDVWVGADTDYADGSDKVVEVAEEVAEVEE 240
 Db 181 GVEFVCCPLAESDNVDSADAEDSDVWVGADTDYADGSDKVVEVAEEVAEVEE 240

```

QY 241 EADDDDDGDEVEEAEPEEYEEATERTTSTATTTTTTSESVEEVVVCSEQAETGPC 300
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 EADDDDEDGDEVEEAEPEEYEEATEKTTSTATTTTTTSESVEEVV- 288
QY 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNPDTEYCMVCGSAMQSLLKTTQEPGLARD 360
Db 289 ----- 288
QY 361 PVKLPPTAASTDPAVDKYLETPGDENEHAHFOKAKERLEAKHRERMSQVMREWEAEARQA 420
    :|||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 289 ---VPTTAASTDPAVDKYLETPGDENEHAHFOKAKERLEAKHRERMSQVMREWEAEARQA 345
QY 421 KNLPKADKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 346 KNLPKADKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 405
QY 481 QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKAAQIRSQVMTLRLVIYER 540
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 406 QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKAAQIRSQVMTLRLVIYER 465
QY 541 MNQSLSLYNVPAVEEIQDEVELLQKEQNSDDVLANNMISEPRISYGNDAIMPSTLET 600
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 466 MNQSLSLYNVPAVEEIQDEVELLQKEQNSDDVLANNMISEPRISYGNDAIMPSTLET 525
QY 601 KTTVELLPVNGEFLSDDLQPHWHSFGADSVPAANTENEVEPVDARPAADRGLTTRPGSGLTN 660
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 526 KTTVELLPVNGEFLSDDLQPHWHPFGVDSVPANTENEVEPVDARPAADRGLTTRPGSGLTN 585
QY 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVGGVVIATVITL 720
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 586 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVGGVVIATVITL 645
QY 721 VMLKKKQYTSIIHGVEVDAAVTPPERHLSKMQQNGYENPTYKFEQMQN 770
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 646 VMLKKKQYTSIIHGVEVDAAVTPPERHLSKMQQNGYENPTYKFEQMQN 695

RESULT 5
P97487
ID P97487 PRELIMINARY; PRT; 695 AA.
AC P97487; P97942;
DT 01-MAY-1997 (TremBLrel. 03, Created)
DT 01-MAY-1997 (TremBLrel. 03, Last sequence update)
DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
DE Hippocampal anlyoid protein.
GN APP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SAMP8; TISSUE=HIPPOCAMPUS;
RA Flood J.F., Kumar V.B., Sasser T., Word I., Morley J.E.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 581-562 FROM N.A.
RC STRAIN=1295V;
RA Wragg M.A., Busfield F., Duff K., Korenblat K., Capecechi M.,
RA Loring J.F., Goate A.M.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U84012; AAB41502.1; -.
DR EMBL; U82624; AAB40919.1; -.
DR HSSP; P05067; 1MRP.
DR MGD; MGI:88059; App.
DR InterPro; IPR001868; A4_APP.
DR InterPro; IPR001255; Beta-APP.
DR Pfam; PF02177; A4_EXTRA; 1.
DR Pfam; PF03494; Beta-APP; 1.
DR PRINTS; PR00203; AMYLOIDA4.
DR SMART; SM00005; A4_EXTRA; 1.
DR PROSITE; P500319; A4_EXTRA; 1.

```

```

DR PROSITE; P500320; A4_INTRA; 1.
SQ SEQUENCE 695 AA; 78414 MW; 9A5FBE2ED261236E CRC64;

Query Match      86.0%; Score 3489.5; DB 11; Length 695;
Best Local Similarity 87.7%; Pred.No. 2.3e-202;
Matches 675; Conservative 6; Mismatches 14; Indels 75; Gaps 1;

QY 1 MLPGLALLLAAMTARALEVPTDGNAGLLAEPOIAFACGLRNHMNVQNGKWDSPSGTK 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 MLPGLALLLAAMTARALEVPTDGNAGLLAEPOIAFACGLRNHMNVQNGKWDSPSGTK 60
QY 61 TCIDTKEGIIQYCOEYVPELQITNVVEANQPVTIQNWKCKGRKQCKTHPHFVTPYCLVG 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 TCIDTKEGIIQYCOEYVPELQITNVVEANQPVTIQNWKCKGRKQCKTHPHFVTPYCLVG 120
QY 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHTVAKETCSKSTNLHDYGMLLPCGIDKFR 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHTVAKETCSKSTNLHDYGMLLPCGIDKFR 180
QY 181 GFVFCVCCPLAEESDVSADAEEDSDVMWGGADTDYADGSEDKVVEAEVEAEVEE 240
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 GFVFCVCCPLAEESDVSADAEEDSDVMWGGADTDYADGSEDKVVEAEVEAEVEE 240
QY 241 EADDDDEDGDEVEEAEPEEYEEATERTTSTATTTTTTSESVEEVVVCSEQAETGPC 300
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 EADDDDEDGDEVEEAEPEEYEEATERTTSTATTTTTTSESVEEVV- 288
QY 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNPDTEYCMVCGSAMQSLLKTTQEPGLARD 360
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 289 ----- 288
QY 361 PVKLPPTAASTDPAVDKYLETPGDENEHAHFOKAKERLEAKHRERMSQVMREWEAEARQA 420
    :|||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 289 ---VPTTAASTDPAADKYLET PGDENEHAHFOKAKERLEAKHRERMSQVMREWEAEARQA 345
QY 421 KNLPKADKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 346 KNLPKADKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 405
QY 481 QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKAAQIRSQVMTLRLVIYER 540
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 406 QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKAAQIRSQVMTLRLVIYER 465
QY 541 MNQSLSLYNVPAVEEIQDEVELLQKEQNSDDVLANNMISEPRISYGNDAIMPSTLET 600
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 466 MNQSLSLYNVPAVEEIQDEVELLQKEQNSDDVLANNMISEPRISYGNDAIMPSTLET 525
QY 601 KTTVELLPVNGEFLSDDLQPHWHSFGADSVPAANTENEVEPVDARPAADRGLTTRPGSGLTN 660
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 526 KTTVELLPVNGEFLSDDLQPHWHPFGVDSVPANTENEVEPVDARPAADRGLTTRPGSGLTN 585
QY 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVGGVVIATVITL 720
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 586 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVGGVVIATVITL 645
QY 721 VMLKKKQYTSIIHGVEVDAAVTPPERHLSKMQQNGYENPTYKFEQMQN 770
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 646 VMLKKKQYTSIIHGVEVDAAVTPPERHLSKMQQNGYENPTYKFEQMQN 695

RESULT 6
Q91963
ID Q91963 PRELIMINARY; PRT; 747 AA.
AC Q91963;
DT 01-NOV-1996 (TremBLrel. 01, Created)
DT 01-NOV-1996 (TremBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
DE APP747.
GN APP747.
OS Xenopus.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Piploidea; Pipidae;
OC Xenopodinae.

```



```

OX NCBI_TaxID=8353;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93129227; PubMed=1282805;
RA Okado H., Okamoto H.;
RT "A Xenopus homolog of the human beta-amyloid precursor protein:
RL development regulation of its gene expression.";
RL Biochem. Biophys. Res. Commun. 189:1561-1568(1992).
DR EMBL: S52417; AAB24853.1; -.
DR HSSP: P05067; 1H23.
DR InterPro: IPR001868; A4_APP.
DR InterPro: IPR001255; Beta_APP.
DR InterPro: IPR002223; Kunitz_BPTI.
DR Pfam: PF02177; A4_EXTRA; 1.
DR Pfam: PF03494; Beta_APP; 1.
DR Pfam: PF00014; Kunitz_BPTI; 1.
DR PRINTS: PR00203; AMYLOIDA4.
DR PRINTS: PR00759; BASICPTASE.
DR ProDom: PD000222; Kunitz_BPTI; 1.
DR SMART: SM00006; A4_EXTRA; 1.
DR SMART: SM00131; KU; 1.
DR PROSITE: PS00319; A4_EXTRA; 1.
DR PROSITE: PS00320; A4_INTRA; 1.
DR PROSITE: PS00280; BPTI_KUNITZ_1; 1.
DR PROSITE: PS00279; BPTI_KUNITZ_2; 1.
KW Serine protease inhibitor.
SQ SEQUENCE 747 AA; 84893 MW; A75E81885681D948 CRC64;

Query Match      83.9%; Score 3403.5; DB 13; Length 747;
Best Local Similarity 85.2%; Pred. No. 3.9e-197;
Matches 645; Conservative 40; Mismatches 45; Indels 27; Gaps 5;

QY 17 ALEVPDGNAGLAEPQIAMF-CGRLENNHNMVQNGKWDSPSGTKCIDTKEGILQYCOE 75
DB 15 ALELVGNGGLAEQIAEFVARLNHNMVQNGKWEVDVSG---CIGTKEGILQYCOE 71

QY 76 VYPELQITNVVEANQPVTTIQNWKGRKQCKTHPHFVPIVYRCVLGFEVSDALLVPDKCF 135
DB 72 VYPELQITNVVEANQPVTTIQNWKCKGRKQCKSTHIVVYRCVLGFEVSDALLVPDKCF 131

QY 136 LHQERMDVCTHLHWHYVAKETSEKSTNLHDYGMLLPGCIDKFRGVEFVCCPLAESDN 195
DB 132 LHQERMDICETHLHWHYVAKESKSMLEHYGMLLPGCIDKFRGVEFVCCPSABES 191

QY 196 VDSADAEEDSDYVWGADTDYADGSEDKVVEVA---EEVEAEVEEEDDEDDEDGE 253
DB 192 FDSADAEEDCDYVWGADADYVDRSDDRAVEAQPDVEEVEVEEEDDEDDEDGE 249

QY 254 VEAEAEPEYBEATERTTSIATTTTTTESVEEYVREVCSEQAETGPCRAMISRWFYDVT 313
DB 250 AEAEPEPEYBEATERTTSIATTTTTTESVEEYVREVCSEQAETGPCRAMISRWFYDVT 309

QY 314 GKCAPEFYGGCGNRNFDTEECYMAVCGSAMQSLLKTTQEPPLARDPVKLPTTAASPD 373
DB 310 SKCAQFYGGCGNRNFDSDCYMAVCGSV-----IPATAASPD 350

QY 374 AVDKYLETGPDENEHAFKAKERLEAKHREMSQVMEWEAEAROKNLPKADKAVIQ 433
DB 351 AVDKYLENPNDENEDRFLKAKERLSEKREKSEYVKEWEAEAROKNLPKADKAVIQ 410

QY 434 HFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITALQAVPPRPRHVFNM 493
DB 411 HFQEKVESLEQEAQKQQLVETHMARVEAMLNDRRLALENYITALQADPPRPRHVFNM 470

QY 494 LKKYVRAEQKDRQHTLKHFEHVRMVDPKAAQIRSQVWTHLRVIVERMNQSLSLYNVPA 553
DB 471 LKKYVRAEQKDRQHTLKHFEHVRMVDPKAAQIRSQVWTHLRVIVERMNQSFSLYKVP 530

QY 554 VAEIQDEVDLQKQKONYSDVLANWISSEPRISYNDALMPSLTETKTVELLPVNGEF 613
DB 531 VAEIQDEVDLQKQKONYSDVLANWISSEPRISYNDALMPSLTETKTVELLPVNGEF 590

QY 614 SLDDLQPHWISFGADSVAPANTENEVEPVDARPAADRLGTTTPGSGGLTNKTEISEYKMDA 673

```

```

Db 591 NIEDLQPHWISFGADSVAPANTENEVEPVDARPAADRLGTTTPGSGGLTNKTEISEYKMD 650
QY 674 EFRHDSGVEVHHQKLVFFAEVGSNGKGAIGLVGGVVIATVITVLMLKKKQVTSIHH 733
DB 651 EYRHTAYEVHHQKLVFFAEVGSNGKGAIGLVGGVVIATVITVLMLKKKQVTSIHH 710
QY 734 GYVEVDAAVTPPEHLSKMQQNGYENPYKPFQEQMN 770
DB 711 GYVEVDAAVTPPEHLSKMQQNGYENPYKPFQEQMN 747

RESULT 7
QDGGJ8
ID Q9DGJ8 PRELIMINARY; PRT; 695 AA.
AC Q9DGJ8;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Beta-amyloid precursor protein 695 isoform.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Sarasa M., Rodolosse A., Sorribas V.;
RT "Cloning of full-length chicken beta-amyloid precursor protein
RT isoforms.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF289218; AAG00593.1; -.
DR HSSP: P05067; 1BA4.
DR InterPro: IPR001868; A4_APP.
DR InterPro: IPR001255; Beta_APP.
DR Pfam: PF02177; A4_EXTRA; 1.
DR Pfam: PF03494; Beta_APP; 1.
DR PRINTS: PR00203; AMYLOIDA4.
DR SMART: SM00006; A4_EXTRA; 1.
DR PROSITE: PS00319; A4_EXTRA; 1.
DR PROSITE: PS00320; A4_INTRA; 1.
SQ SEQUENCE 695 AA; 78565 MW; F201ED02AEC86D95 CRC64;

Query Match      83.2%; Score 3377.5; DB 13; Length 695;
Best Local Similarity 84.7%; Pred. No. 1.3e-195;
Matches 654; Conservative 18; Mismatches 21; Indels 79; Gaps 4;

QY 1 MLPLGLALLLLAATAWTALEVPDGNAGLLAEFQIAMFCGRLENNHNMVQNGKWDSPSGTK 60
DB 1 MLPLGLALLLLAAGAAARALEVPADGNAGLLAEFQIAMFCGRLENNHNMVQNGKWDSPSGTK 60

QY 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTTIQNWKGRKQCKTHPHFVPIVYRCVLG 120
DB 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTTIQNWKGRKQCKTHPHFVPIVYRCVLG 120

QY 121 EYVSDALLVPDKCFKHQERMDVCTHLHWHYVAKETSEKSTNLHDYGMLLPGCIDKFR 180
DB 121 EYVSDALLVPDKCFKHQERMDVCTHLHWHYVAKESKSMLEHYGMLLPGCIDKFR 180

QY 181 GVEFYCCPLAESDNVDSADAEEDSDYVWGADTDYADGSEDKVVE--VAEEVEAEVE 238
DB 181 GVEFYCCPLAESDNVDSADAEEDSDYVWGADADYADGSDDKVVEEQDEDELTVVE 240

QY 239 EEEADDDDEDGDDVVEEAEPEEATERTTSIATTTTTTESVEEYVREVCSEQAETG 298
DB 241 EDADADD--DDDDGEI--BETEEYBEATERTTSIATTTTTTESVEEYVREVCSEQAETG 288

QY 299 PCRAMISRWFYDVTBQKCAPFFYGGCGNRNFDTEECYMAVCGSAMQSLLKTTQEPPLA 358
DB 289 ----- 288

QY 359 RDPVKLPTTAASPDVADKYLETGPDENEHAFKAKERLEAKHREMSQVMEWEAEAR 418

```

```

Db 289 -----VPTTAASTPDAVDKYLETPGDENEHAHFQAKERLEAKHRERMSQVWREWEAEER 343
QY 419 QAKNLPKADKKAVIQHFQEKVESLEGEAAANERQQLVETHMARVEAMLNDRRLALENYIT 478
Db 344 QAKNLPKADKKAVIQHFQEKVESLEGEAAANERQQLVETHMARVEAMLNDRRLALENYIT 403
QY 479 ALQAVPPRPRHVNMLKKYVRAEQKDRQHTLKFHFHVRVDPKKAQIRSOVMTHLRVIY 538
Db 404 ALQAVPPRPRHVNMLKKYVRAEQKDRQHTLKFHFHVRVDPKKAQIRSOVMTHLRVIY 463
QY 539 ERMNQSLLYNPVAVEEIQDEVDLLOKEQNYSDVLANMISEPRIISYGNDAIMPST 598
Db 464 ERMNQSLLYNPVAVEEIQDEVDLLOKEQNSDVLANMISEPRIISYGNDAIMPST 523
QY 599 ETKTTVELLPVNGEFLDQLQPHWSEFADSVDPANTENEVEPVDARPAADRGTLTRPGSL 658
Db 524 ETKTTVELLPVNGEFLDQLQPHWSEFADSVDPANTENEVEPVDARPAADRGTLTRPGSL 583
QY 659 TNKTEEISEVKMDAEFRHDSGEYVHHQKLVFPAEDVGSNGKAGIIGLMVGGVVIATVIVI 718
Db 584 TNKTEEISEVKMDAEFRHDSGEYVHHQKLVFPAEDVGSNGKAGIIGLMVGGVVIATVIVI 643
QY 719 TLVMLKKKQYTSIHGGVVEVDAATPEERHLKSKMQONGYENPTYKFFEQMON 770
Db 644 TLVMLKKKQYTSIHGGVVEVDAATPEERHLKSKMQONGYENPTYKFFEQMON 695

RESULT 8
Q98SGO PRELIMINARY; PRT; 693 AA.
AC Q98SGO;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Beta-amyloid precursor protein A.
GN APP.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OC NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Van den Hurk W.H.;
RL Thesis (2001), Department of Biological Sciences,
RL University of Nijmegen, Nijmegen, Netherlands.
DR EMBL; AJ298150; CAC37193.1;
DR HSP; P05067; 1H23.
DR InterPro; IPR001868; A4_APP.
DR InterPro; IPR001255; Beta-APP.
DR Pfam; PF02177; A4_EXTRA; 1.
DR Pfam; PF03494; Beta-APP; 1.
DR PRINTS; PR00203; AMYLOIDA4.
DR SMART; SM00006; A4_EXTRA; 1.
DR PROSITE; PS00319; A4_EXTRA; 1.
DR PROSITE; PS00320; A4_INTRA; 1.
FT SIGNAL.
KW Signal.
SQ SEQUENCE 693 AA; 78568 MW; CAFIDF655CIAB653 CRC64;

Query Match 78.0%; Score 3163.5; DB 13; Length 693;
Best Local Similarity 79.1%; Pred. No. 1.1e-182;
Matches 611; Conservative 37; Mismatches 43; Indels 81; Gaps 5;

QY 1 MLPGLALLAAWTARALEVPTDGNAGLLAEPOIAMFCGLNMHMNVQNGKWDSPSGTK 60
Db 1 MLPHTILLVITV-GALALEVPADNGGLLAEPOIAMFCGLNMHMNVQNGKWDSPSGTK 59
QY 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTIQNCKRGRQCKTHPHFVIPYRCLVG 120
Db 60 GCIGTKEGILQYCOEYVPELQITNVVEANQPVTIQNCKRGRQCKSRTHVIVPYRCLVG 119
QY 121 EFVSQDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCEKSTNLHDYGMLLPCGIDKFR 180

```

```

Db 120 EFVSQDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCEKSTNLHDYGMLLPCGIDKFR 179
QY 181 GVEFFCCPLABESQNVSDADAEEDSDVWGGADTDYADGSEDKVVEYA--EEBEVAEVE 238
Db 180 GVEFFCCPLABESQNVSDADAEEDSDVWGGADTDYADGSEDKVVEYA--EEBEVAEVE 238
QY 239 EEEADDDDEDDGDEVEEAEPEEATERTSTIATTTTTTSTESVEEVVREVCSEQAETG 298
Db 239 EEEADDDDEDDGDEVEEAEPEEATERTSTIATTTTTTSTESVEEVVREVCSEQAETG 286
QY 239 PCRAMISWYFDVTGKCAPFFYGGCGGNRNFDTEECVMAVCGSAMQSLLKTTQBPLA 358
Db 287 ----- 286
QY 359 RDPVKLPPTAASTPDAVDKYLETPGDENEHAHFQAKERLEAKHRERMSQVWREWEAEER 418
Db 287 -----VPATAASTPDAVDKYLETPGDENEHAHFQAKERLEAKHRERMSQVWREWEAEER 341
QY 419 QAKNLPKADKKAVIQHFQEKVESLEGEAAANERQQLVETHMARVEAMLNDRRLALENYIT 478
Db 342 QAKNLPKADKKAVIQHFQEKVESLEGEAAANERQQLVETHMARVEAMLNDRRLALENYIT 401
QY 479 ALQAVPPRPRHVNMLKKYVRAEQKDRQHTLKFHFHVRVDPKKAQIRSOVMTHLRVIY 538
Db 402 ALQAVPPRPRHVNMLKKYVRAEQKDRQHTLKFHFHVRVDPKKAQIRSOVMTHLRVIN 461
QY 539 ERMNQSLLYNPVAVEEIQDEVDLLOKEQNYSDVLANMISEPRIISYGNDAIMPST 598
Db 462 ERMNQSLLYNPVAVEEIQDEVDLLOKEQNYSDVLANMISEPRIISYGNDAIMPST 521
QY 599 ETKTTVELLPVNGEFLDQLQPHWSEFADSVDPANTENEVEPVDARPAADRGTLTRPGSL 658
Db 522 ETKTTVELLPVNGEFLDQLQPHWSEFADSVDPANTENEVEPVDARPAADRGTLTRPGSL 581
QY 659 TNKTEEISEVKMDAEFRHDSGEYVHHQKLVFPAEDVGSNGKAGIIGLMVGGVVIATVIVI 718
Db 582 TNKTEEISEVKMDAEFRHDSGEYVHHQKLVFPAEDVGSNGKAGIIGLMVGGVVIATVIVI 641
QY 719 TLVMLKKKQYTSIHGGVVEVDAATPEERHLKSKMQONGYENPTYKFFEQMON 770
Db 642 TLVMLKKKQYTSIHGGVVEVDAATPEERHLKSKMQONGYENPTYKFFEQMON 693

RESULT 9
Q98SF9 PRELIMINARY; PRT; 695 AA.
AC Q98SF9;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Beta-amyloid precursor protein B.
GN APP.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OC NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Van den Hurk W.H.;
RL Thesis (2001), Department of Biological Sciences,
RL University of Nijmegen, Nijmegen, Netherlands.
DR EMBL; AJ298151; CAC37194.1;
DR HSP; P05067; 1H23.
DR InterPro; IPR001868; A4_APP.
DR InterPro; IPR001255; Beta-APP.
DR Pfam; PF02177; A4_EXTRA; 1.
DR Pfam; PF03494; Beta-APP; 1.
DR PRINTS; PR00203; AMYLOIDA4.
DR SMART; SM00006; A4_EXTRA; 1.
DR PROSITE; PS00319; A4_EXTRA; 1.
DR PROSITE; PS00320; A4_INTRA; 1.

```

KW Signal.
 FT SIGNAL
 SQ SEQUENCE 695 AA; 78803 MW; DC14EB02AFB0204A CRC64;

Query Match 77.7%; Score 3152.5; DB 13; Length 695;
 Best Local Similarity 78.9%; Pred. No. 4.9e-182;
 Matches 609; Conservative 39; Mismatches 45; Indels 79; Gaps 5;

QY 1 MLPGALLLAATAARALEVPTDGNAGLLAEPOIAFCGRLNNHNVQNGKWDSPSGTK 60
 DB 1 MLPHITLLVLA-GALALEVPADGNGLLAEPOIAFCGRLNNHNVQNGKWDSPSGTK 59

QY 61 TCIDTREGILQCOEYVPELQITNVVEANQPTIQNCKKGRKQCKTHPHFVPIYRCLVG 120
 DB 60 CGTGTREGILQCOEYVPELQITNVVEANQPTIQNCKKGRKQCKTHPHFVPIYRCLVG 119

QY 121 EFVSDALLVPDKCKFHQERMDVCETHLHWHHTVAKETSEKSTNLDHYGMLLPGCIDKFR 180
 DB 120 EFVSDALLVPDKCKFHQERMDVCETHLHWHHTVAKETSEKSTNLDHYGMLLPGCIDKFR 179

QY 181 GVEFVCCPLAESDNVDSADAEEDSDVMWGGADTDYADGSEKVVVEV-AEEVEEVAEVE 238
 DB 180 GVEFVCCPLAESDNVDSADAEEDSDVMWGGADTDYADGSEKVVVEV-AEEVEEVAEVE 238

QY 239 EEEADDEDEDCGDEVEEAEPEATERTTSIATTTTTTSTESVEEYVVEVCSEQAETG 298
 DB 239 EEEADDEDEDCGDEVEEAEPEATERTTSIATTTTTTSTESVEEYVVEVCSEQAETG 298

QY 299 PCRAMISRWYFDVTEGKCAPFFYGGCGGNRNDFEEYCMVCGSAMQSLLKTTQEPPLA 358
 DB 290 ----- 289

QY 359 RDPVKLPTTAASTPDADVKYLETTPGDENEHAFQKAKERLEAKHRMSOVMWEWEAEAR 418
 DB 290 -----PATAVSTPDADVKYLETTPGDENEHAFQKAKERLEAKHRMSOVMWEWEAEAR 418

QY 419 QAKNLPKADKAVIOHFQEKVESLEQEAANERQQLVETHMARVAMLNDRRLALENYIT 478
 DB 344 QAKNLPKADKAVIOHFQEKVESLEQEAANERQQLVETHMARVAMLNDRRLALENYIT 403

QY 479 ALQAVPPRRHVNFMKKYVRAEQKDRQHTLKHFHVRMVDPKAAQIRSOVMTHLRVY 538
 DB 404 ALQAVPPRRHVNFMKKYVRAEQKDRQHTLKHFHVRMVDPKAAQIRSOVMTHLRVY 463

QY 539 ERMNOSLLYNVPAVEEIQDEVDLQKQNYSDVLANMISEPRISYGNDAIMPST 598
 DB 464 ERMNOSLLYNVPAVEEIQDEVDLQKQNYSDVLANMISEPRISYGNDAIMPST 523

QY 599 ETKTTVELLPVNGEFLDLPWHSFGADSVPAANTEVEPEVDARPAADRGLTTRPGSL 658
 DB 524 ETKTTVELLPVNGEFLDLPWHSFGADSVPAANTEVEPEVDARPAADRGLTTRPGSL 583

QY 659 TNIKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEVDYGSNKGAIIGLMVGGVVIATVVI 718
 DB 584 TNIKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEVDYGSNKGAIIGLMVGGVVIATVVI 643

QY 719 TLVMLKKKQYTSIHGGVVEVDAATPEERHLSKMQQNGYENPTYKFFEQM 770
 DB 644 TLVMLKKKQYTSIHGGVVEVDAATPEERHLSKMQQNGYENPTYKFFEQM 695

RESULT 10
 Q99K32 ID Q99K32 PRELIMINARY; PRT; 607 AA.
 AC Q99K32;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Hypothetical 68.4 kDa protein (Fragment).
 OS App.
 OS Mus musculus (House).
 OC Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC005490; AAH05490.1; -;
 DR HSSP; P05067; IAAP.
 DR MGD; MGI:88059; App.
 DR InterPro; IPR001868; A4_APP.
 DR InterPro; IPR001255; Beta-APP.
 DR InterPro; IPR002223; Kunitz_BPTI.
 DR Pfam; PF02177; A4_EXTRA; 1.
 DR Pfam; PF03494; Beta-APP; 1.
 DR Pfam; PF00014; Kunitz_BPTI; 1.
 DR PRINTS; PR00203; AMYLOIDA4.
 DR PRINTS; PR00759; BASICPTASE.
 DR ProDom; PD000222; Kunitz_BPTI; 1.
 DR SMART; SM00131; KU; 1.
 DR PROSITE; PS00319; A4_EXTRA; 1.
 DR PROSITE; PS00320; A4_INTRA; 1.
 DR PROSITE; PS00280; BPTI_KUNITZ_1; 1.
 DR PROSITE; PS00279; BPTI_KUNITZ_2; 1.
 KW Hypothetical protein; Serine protease inhibitor.
 FT NON_TER 1
 SQ SEQUENCE 607 AA; 68391 MW; BF802214CBA7D172 CRC64;

Query Match 75.5%; Score 3065; DB 11; Length 607;
 Best Local Similarity 96.9%; Pred. No. 7.7e-177;
 Matches 588; Conservative 5; Mismatches 14; Indels 0; Gaps 0;

QY 164 NLHDYGMLLPCGIDKFRGVEFVCCPLAESDNVDSADAEEDSDVMWGGADTDYADGSE 223
 DB 1 NLHDYGMLLPCGIDKFRGVEFVCCPLAESDNVDSADAEEDSDVMWGGADTDYADGSE 60

QY 224 KYVEAAEEVEAEVEEEDDEDCGDEVEEAEPEATERTTSIATTTTTTSTESV 283
 DB 61 KYVEAAEEVEAEVEEEDDEDCGDEVEEAEPEATERTTSIATTTTTTSTESV 120

QY 284 EEWVEVCSEQAETGPCRAMISRWYFDVTEGKCAPFFYGGCGGNRNDFEEYCMVCGS 343
 DB 121 EEWVEVCSEQAETGPCRAMISRWYFDVTEGKCAPFFYGGCGGNRNDFEEYCMVCGS 180

QY 344 AMSQLLKTTOEPLARDPVKLTPTTAASTPDADVKYLETTPGDENEHAFQKAKERLEAKHR 403
 DB 181 VSTQSLKTTSEPLPDQDPDKLPTTAASTPDADVKYLETTPGDENEHAFQKAKERLEAKHR 240

QY 404 ERMSQVMEWEAEAEQAANLPKADKAVIOHFQEKVESLEQEAANERQQLVETHMARVEA 463
 DB 241 ERMSQVMEWEAEAEQAANLPKADKAVIOHFQEKVESLEQEAANERQQLVETHMARVEA 300

QY 464 MLNDRRLALENYITLQAVPPRRHVNFMKKYVRAEQKDRQHTLKHFHVRMVDPKKA 523
 DB 301 MLNDRRLALENYITLQAVPPRRHVNFMKKYVRAEQKDRQHTLKHFHVRMVDPKKA 360

QY 524 AQIRSQVMTHLRVIERMNSLSLLYNVPAVEEIQDEVDLQKQNYSDVLANMISE 583
 DB 361 AQIRSQVMTHLRVIERMNSLSLLYNVPAVEEIQDEVDLQKQNYSDVLANMISE 420

QY 584 PRISYGNDAIMPSTETKTTVELLPVNGEFLDLPWHSFGADSVPAANTEVEPEVDAR 643
 DB 421 PRISYGNDAIMPSTETKTTVELLPVNGEFLDLPWHSFGADSVPAANTEVEPEVDAR 480

QY 644 PAADRGLTTRPGSLTNIKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEVDYGSNKGAI 703
 DB 481 PAADRGLTTRPGSLTNIKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEVDYGSNKGAI 540

QY 704 GLMVGGVVIATVITLVMLKKKQYTSIHGGVVEVDAATPEERHLSKMQQNGYENPTYK 763
 DB 541 GLMVGGVVIATVITLVMLKKKQYTSIHGGVVEVDAATPEERHLSKMQQNGYENPTYK 600

QY 764 FFEQMQN 770
 DB 601 FFEQMQN 607

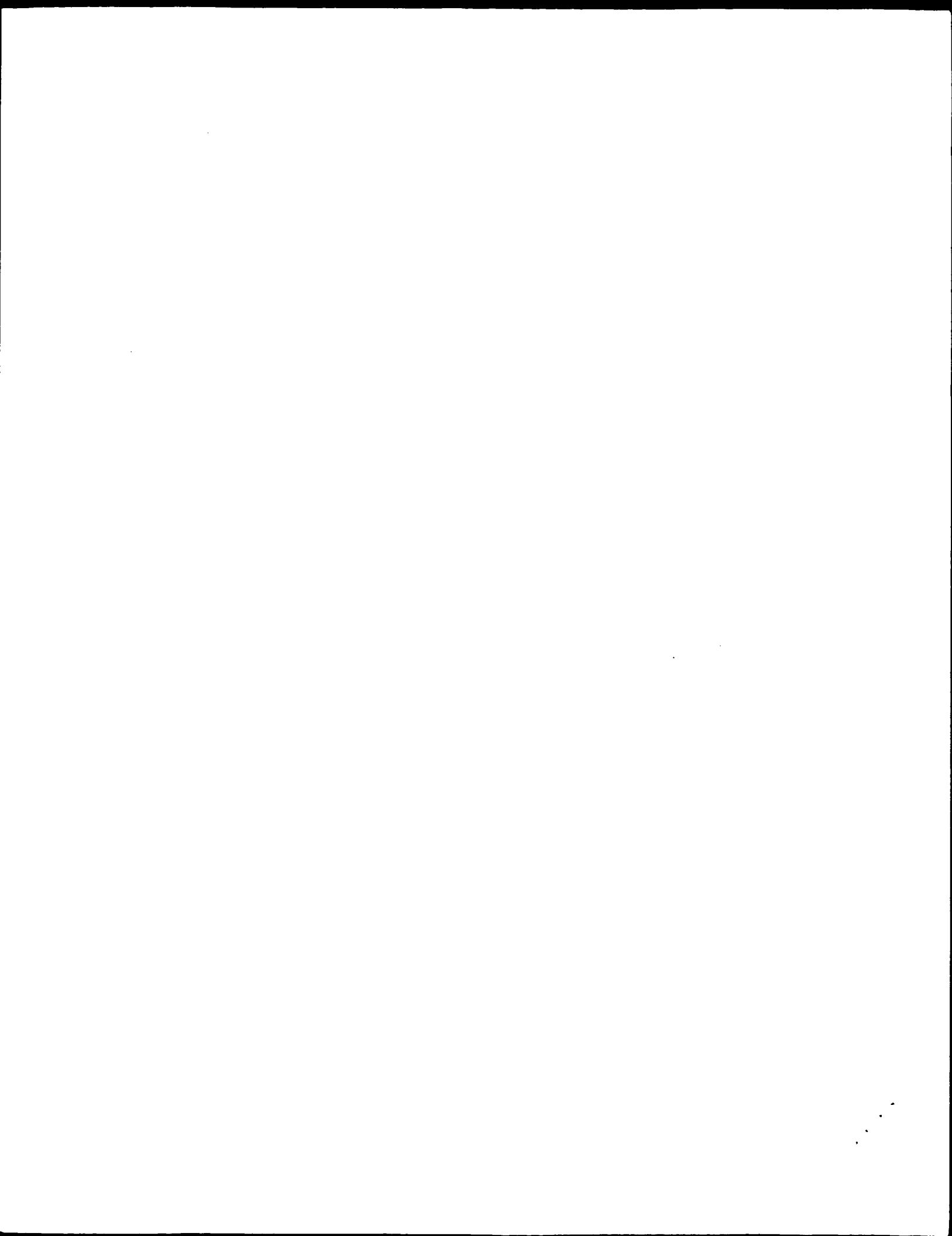
RESULT 11	
Q57394	O57394 PRELIMINARY; PRT; 699 AA.
AC	O57394; 1998 (Tremblrel. 06, Created)
DT	01-JUN-1998 (Tremblrel. 06, Last sequence update)
DT	01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE	EL amyloid precursor protein 699.
GN	EL App699.
OS	Narke japonica (Electric ray).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC	Elasmobranchii; Squala; Hymnosquala; Pristionotidae; Batoidae;
OC	Torpediniformes; Narcinoidae; Narkidae; Narke.
OX	NCBI_TaxID=62965;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE-ELECTRIC LOBE;
RX	MEDLINE=98129705; PubMed=9461486;
RA	Iijima K., Lee D.-S., Okutsu J., Tomita S., Hirashima N., Kirino Y.,
RA	Suzuki T.;
RT	"cDNA isolation of Alzheimer's amyloid precursor protein from
RT	cholinergic nerve terminals of the electric organ of the electric
RT	ray."
RL	Biochem. J. 330:29-33(1998).
DR	EMBL; AB005544; BAA24230.1; -.
DR	HSSP; P05067; 1H23.
DR	InterPro; IPR001868; A4_APP.
DR	InterPro; IPR001255; Beta-APP.
DR	Pfam; PF02177; A4_EXTRA; 1.
DR	Pfam; PF03494; Beta-APP; 1.
DR	PRINTS; PR00203; AMYLOIDA4.
DR	SMART; SM00006; A4_EXTRA; 1.
DR	PROSITE; PS00319; A4_EXTRA; 1.
DR	PROSITE; PS00320; A4_INTRA; 1.
SQ	SEQUENCE 699 AA; 78879 MW; 952915C309D50E5C CRC64;
Query Match 71.8%; Score 2914; DB 13; Length 699;	
Best Local Similarity 72.9%; Pred. No. 1.2e-167;	
Matches 568; Conservative 59; Mismatches 58; Indels 94; Gaps 9;	
QY	2 LPG-LALLLAAWTA-----RALEVPTDGNAGLL-AEQIAMFCGRNLNMHNVQNGKW 52
DB	1 LG-LGMLLLAAALAAVLAPLCRALEVPTDGGAGLAAEQIAMFCGRNLNMHNVQNGKW 64
QY	53 DSDPSGKTKCIDTKGILQYQEVYPELQITNVVEANQPTIQNWCKRGRKCKTHPHFV 112
DB	65 VSDPSGNTNCTCGTKEGILRYQEVYPDQITNVVEANQPTIQNWCKRGRKCKGHPHIV 124
QY	113 IPYRCLVGFVSDALLVPDKCKFLHQRMDVCEETHLWHTVAKETCSKSTNLHDYGMLL 172
DB	125 VYRCLVGFVSDALLVPDKCKFLHQRMDVCEETHLWHTVAKETCGDKINLHDYGMLL 184
QY	173 PGCIDKFRGVFVCCPLAEESDNVSDAAEEDSDVWNGGADTDVADGSEDKVVEAE 232
DB	185 PCGIDFRGVFVCCPIPENDKIDS-DMDEEDSDVWNGGADTDVADG-DKTV---EE 238
QY	233 EVAEVEEEDDE 291
DB	239 KPTEEEDE 292
QY	292 SEQAEATGPKRAMISRWFVDVTEGKCAPFFYGGCGGNRNFDEEYCMVCGSAMSQSLK 351
DB	293 ----- 292
QY	352 TTOEPLARDPVKLPTTAATPDVNDKYLETPGDENEHAHFQAKERLEAKHRERMSQVNR 411
DB	293 -----VPTTAATPDVNDKYLETPGDENEHAHFQAKERLEAKHRERMSKIMR 340
QY	412 EWEAEARQAKNLPKADKAVIQHFQKESLQEAANERQQLVETVMARVEAMLNDRRL 471
DB	341 EWEAEARQAKNLPKADKAVIQRFQENVESLQEAASERQQLVETVMARVEAMLNDRRI 400

188 CP-AEARDMDSTKEDADDSDVWVGADNDYSDNSVMYREPEPAEQQEQETRPSVVEEEEG 246
233 EVAVEEBEE-----ADDEDDEGDVEEAEAEPEYEATERTTTIA 273
247 EVAQEDDEEEVLDTDQDGDGEDHEAADDEEEEDVDEIDAFGESDDVDADEPTTNA 306
274 ---TTTTTTTSEBVREYVCSQAETGPCRAMISRMYFDVTGKCA-PFFYGCNGNRN 329
307 MTTTTTTTTTSEBVREYVCMFCWAHADGPGCTASPSWSYFDAVDGRTMYELMYGCGCGNMN 366
330 NFDTEECYMCVGSAMSQSLLKTQEPFLARDPVKLPPTTAASTPDADVXYLETPGDNEHA 399
367 NFEESEYCLSYCSSV-----VPTDMPSSPDADVHYLETPADNEHA 407
390 HFQAKERLEAKHRMSQVWREWEAEARQAKNLPKADKKAVIOHFOEKVESLEQEAANE 449
408 HFQAKESLAKHRMSQVWREWEAEARQAKNLPRADKKIVIORFOEKVEALEQEAASE 467
450 RQQLVETHMARVEAMLNDRRLALENITALQAVPPRPHRVFNMLKKYVRAEQKDRQHTL 509
468 RQQLVETHMARVEALLNDRRLALENLTALQADPPRPHRVFSLKKYVRAEQKDRQHTL 527
510 KHFPHVRMDPKKAAQIRTSQVMTHLRVYIERMNSQLSLYNYPVAAEIODEVDELQKE 569
528 KHFPHVRMDPKKAAQIRPQVLTHLRVIERMNSQLGLLYKVPGVADDIQDQV-ELLQRE 586
570 QNYSDDVLANNISEPRISYNDALMPSLSTETKITVELLPVNGEESLDDLPWH--SFGAD 627
587 QAEMAQOLANLQTDVRSYNDALMPDQELGDQADLLP--QEDTLGGVGFVHPESFN-- 642
628 SVPANTENEPEVDARPAADRGLTTRPGSLTNIKTEEISEVKWDAEFRHDSGYEVHHQK 687
643 --QLNTENQVEPVDSPRTPFERGVPTP---VTCKSMEANPELSEMETEDRQSTEYEVHHQK 697
688 LVFFAEDVGSNGKGAIGLMYGVGVVIATVITILVMLKKQYTSIIHHGVVEVDAAVTPEER 747
698 LVFFAEDVGSNGKGAIGLMYGVGVVIATVITILVMLKKQYTSIIHHGIIIEVDAAVTPEER 757
748 HLSKMQONGYENPTYKFEQMON 770
758 HLSKMQONGYENPTYKFEQMON 780

RESULT 15
Q9PVL1
ID Q9PVL1 PRELIMINARY; PRT: 569 AA.
AC Q9PVL1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Amyloid protein (Fragment).
GN APP.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopteria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRIN;
RA Coulson E.J., Palliga K., Beyreuther K., Masters C.L.;
RT "What the evolution of the amyloid protein precursor superegene family
RT tells us about its function";
RL Neurochem. Int. 0:0-0(2000).
DR EMBL; AF030341; AAF12698.1; -.
DR HSSP; P05067; 1BA4.
DR InterPro; IPR001868; A4_APP.
DR InterPro; IPR001255; Beta_APP.
DR Pfam; PF02177; A4_EXTPA; 1.
DR Pfam; PF03494; Beta_APP; 1.
DR PRINTS; PR00203; AMYLOIDA4.
DR SMART; SM00006; A4_EXTRA; 1.
DR PROSITE; PS00319; A4_EXTRA; 1.

DR PROSITE; PS00320; A4_INTRA; 1.
FT NON_TER 1
SQ SEQUENCE 569 AA; 64753 MW; 0AB8B851863A19D CRC64;
Query Match 67.0%; Score 2717; DB 13; Length 569;
Best Local Similarity 82.5%; Pred. No. 6.8e-156;
Matches 534; Conservative 15; Mismatches 18; Indels 80; Gaps 5;
QY 126 ALLVPPDKCKFLHGERMDVCEHLLHWHITVAKETCSKSTNLHDYGMLLPGIDIKFRGVFV 185
Db 1 ALLVPPDKCKLLHGERMDVCEHLLHWHITVAKESCKSMNLHDYGMLLSCGIDIKFRGVFV 60
QY 186 CCPLAESONVDSADAEEDSDSVWGGADTDYADGSEDKVV- -VAEEVEVAEVEEAD 243
Db 61 CCPLAESONLDSADAEEDSDSVWGGADADYADGSDDKVVEEQPEDEELIIVVEDEDAD 120
QY 244 DDEDDGDEVEEAEPEEYEEATERTTSIATTTTTTSTESVEEVVREVCSEQAETGPCRAM 303
Db 121 DD-DDDDGDEI-BETEEVEEATERTTSIATTTTTTSTESVEEVV- - 163
QY 304 ISRWYFDVTGKCAPFYGGCGNRNFDTEECYMAVCGSAMQSLSLLKTTQEP LARDPVK 363
Db 164 ----- 163
QY 364 LPTTAASTPDVADKYLETPGDENEHAHFOKAKERLEAKHERMSQVMREWEAEARQAKNL 423
Db 164 VPTTAASTPDVADKYLETPGDENEHAHFOKAKERLEAKHERMSQVMREWEAEARQAKNL 223
QY 424 PKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITALQAV 483
Db 224 PKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITALQTV 283
QY 484 PPRPRHVNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVYIERMNO 543
Db 284 PPRPRHVNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVYIERMNO 343
QY 544 SLSLYNVPAAVEIQDEVDDELQKEQNYSDVLANMISEPRISYGNALMPSLTETKT 603
Db 344 SLSLYNVPAAVEIQDEVDDELQKEQNYSDVLANMISEPRISYGNALMPTLTETKT 403
QY 604 VELLPVNGEFLDQLPWHSGFADSVDPANTENEVEVPDARPAADRGLTTRPGSGLTNKT 663
Db 404 VELLPVNGEFLDQLPWHSPFVDSVPANTENEVEVPDARPAADRGLTTRPGSGLTNKT 463
QY 664 EETSEVKMDAEFRHDSGYEVHOKLVFFAEVDGSKNGAIIGLMVGWVIATVIVITLVML 723
Db 464 EEVSEVKMDAEFRHDSGYEVHOKLVFFAEVDGSKNGAIIGLMVGWVIATVIVITLVML 523
QY 724 KKKQYTSIHGGVVEVDAAYTPPEERHLSKMQQNGYENPTYKFFEQMON 770
Db 524 KKKQYTSIHGGVVEVDAAYTP-ERHLSKMQQNGYENPTYKFFEQMON 569

Search completed: January 28, 2003, 14:18:15
Job time : 42.9101 secs



GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 13:58:08 ; Search time 42.9063 seconds
(without alignments)
2391.326 Million cell updates/sec

Title: US-09-904-987-2

Perfect score: 4058

Sequence: 1 MLPGLALLLLAAWTARALEV.....KMQQNGYENPTYKFEQMQN 770

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002:*
1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4058	100.0	770	10 AAP94775	Novel amyloid prec
2	4058	100.0	770	14 AAR41546	Mutated APP770 exo
3	4058	100.0	770	15 AAR63442	Amyloid protein pr
4	4058	100.0	770	19 AAW40130	Human APP770 prote
5	4058	100.0	770	20 AAW97996	Human amyloid prec
6	4058	100.0	770	22 AAE10648	Human amyloid prot
7	4058	100.0	770	22 AAE11762	Human amyloid prec
8	4058	100.0	770	22 AAE06893	Human amyloid prec
9	4058	100.0	770	22 AAU06522	Human partial Amyl
10	4058	100.0	770	22 AAE02600	Human amyloid prec

11	4058	100.0	770	23 ABB78609	Human APP770 prote
12	4058	100.0	770	23 AAG58317	Human amyloid prec
13	4058	100.0	772	22 AAE10650	Human amyloid prot
14	4058	100.0	772	22 AAE06895	Human amyloid prec
15	4058	100.0	772	22 AAU06624	Human beta-amyloid
16	4058	100.0	772	22 AAU07223	Human amyloid prec
17	4058	100.0	772	22 AAE02602	Human amyloid prec
18	4058	100.0	772	23 ABB78611	Human APP770-KK pr
19	4054	99.9	770	15 AAR62505	Amyloid precursor
20	4053	99.9	770	18 AAW19500	APP770 mutant A-be
21	4053	99.9	770	18 AAW19497	APP770 mutant A-be
22	4053	99.9	770	18 AAW19485	APP770 mutant A-be
23	4053	99.9	770	18 AAW19482	APP770 mutant A-be
24	4053	99.9	770	22 AAE06913	Human amyloid prec
25	4050	99.8	770	22 AAE06912	Human amyloid prec
26	4047	99.7	768	23 AAU08959	Human amyloid beta
27	4047	99.7	770	18 AAW19506	APP770 mutant A-be
28	4047	99.7	770	18 AAW19491	APP770 mutant A-be
29	4046	99.7	770	13 AAR26340	APP770. Homo sapi
30	4046	99.7	770	18 AAW19488	APP770 mutant A-be
31	4046	99.7	770	18 AAW19503	APP770 mutant A-be
32	4039	99.5	770	11 AAR05717	NAP gene product a
33	3943.5	97.2	751	12 AAR10022	Beta-amyloid-relat
34	3943.5	97.2	751	13 AAR20328	Sequence encoded b
35	3943.5	97.2	751	20 AAY08615	Human beta-amyloid
36	3943.5	97.2	751	20 AAY08605	Human beta-amyloid
37	3943.5	97.2	751	22 AAE10649	Human amyloid prot
38	3943.5	97.2	751	22 AAE06894	Human amyloid prec
39	3943.5	97.2	751	22 AAU06623	Human partial Amyl
40	3943.5	97.2	751	22 AAE02601	Human amyloid prec
41	3943.5	97.2	751	23 AAO18050	Amyloid precursor
42	3943.5	97.2	751	23 ABB78610	Human APP751 prote
43	3943.5	97.2	751	23 AAG68316	Human amyloid prec
44	3943.5	97.2	753	22 AAE10651	Human amyloid prot
45	3943.5	97.2	753	22 AAE06896	Human amyloid prec

ALIGNMENTS

RESULT 1

AAP94775

ID AAP94775 standard; protein; 770 AA.

XX AC AAP94775;

XX DT 05-JUL-1990 (first entry)

XX DE Novel amyloid precursor protein (NAP).

XX KW SPAP; amyloid precursor protein; dysbolism; INS76; ds.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Active-site 289...345

FT /*tag= a

FT /*label=INS76

XX PN EP304013-A.

XX PD 22-FEB-1989.

XX PF 16-AUG-1988; 88EP-0113283.

XX PR 15-AUG-1987; 87JJP-0203298.

XX PR 21-AUG-1987; 87JP-0207995.

XX PR 18-NOV-1987; 87JJP-0291404.

XX PR 11-DEC-1987; 87JJP-0313228.

XX PR 05-FEB-1988; 88JJP-0025260.

XX PR 10-FEB-1988; 88JJP-0029366.

XX PR 19-FEB-1988; 88JJP-0037905.

XX PR 25-MAY-1988; 88JJP-0125660.

```

XX PA (ASAH ) ASahi KASEI KOGYO.
XX PI Kitaguchi N, Takahashi Y, Tokushima Y, Itoh H;
XX DR WPI; 1989-055458/08.
XX DR N-PSDB; AAN91049.
XX PT Human senile plaque amyloid precursor protein and DNA -
XX PT used for study and diagnosis of dysbolism in the central nervous
XX PT system.
XX PS Disclosure; ; 108pp; English.
XX CC AA. sequence from 289 to 345 encodes INS76 which is claimed as a senile
XX CC plaque amyloid precursor (SPAP), useful for diagnosis of dysbolism in the
XX CC CNS such as senile dementia.
XX SQ Sequence 770 AA;

Query Match 100.0%; Score 4058; DB 10; Length 770;
Best Local Similarity 100.0%; Pred. No. 5.3e-289;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPOIAMECGRLLNMHMNVONGKWDSPSGTK 60
Db 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPOIAMECGRLLNMHMNVONGKWDSPSGTK 60

Qy 61 TCIDTKEGILQYCEVYPELQITNVVEANQPVTIQNCKRGRKCKTHPHFVPIYRCLVG 120
Db 61 TCIDTKEGILQYCEVYPELQITNVVEANQPVTIQNCKRGRKCKTHPHFVPIYRCLVG 120

Qy 121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
Db 121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180

Qy 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSDKVEVVAEEVAEEVEE 240
Db 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSDKVEVVAEEVAEEVEE 240

Qy 241 EADDEDDEGDEVEEAEPEYBATERTTSIATTTTTTSTESVEEVREVCSQAETGPC 300
Db 241 EADDEDDEGDEVEEAEPEYBATERTTSIATTTTTTSTESVEEVREVCSQAETGPC 300

Qy 301 RAMLSRWYFDVTEGKCAPFFYGGCGGRNPFDEEYCMVCGSAMSGSLKTTQEP LARD 360
Db 301 RAMLSRWYFDVTEGKCAPFFYGGCGGRNPFDEEYCMVCGSAMSGSLKTTQEP LARD 360

Qy 361 PVKLPPTTAASPDADVXYLETDPGDENEHAFQAKERLEAKHRERMSOVNREWEAEERQA 420
Db 361 PVKLPPTTAASPDADVXYLETDPGDENEHAFQAKERLEAKHRERMSOVNREWEAEERQA 420

Qy 421 KNLPAKDKAVIOHFQKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENITAL 480
Db 421 KNLPAKDKAVIOHFQKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENITAL 480

Qy 481 QAVPPRRHVFNMMLKKYVRAEQDRQHTLKHFEHVRMVDPKKAAQTRSQVWTHLRIYER 540
Db 481 QAVPPRRHVFNMMLKKYVRAEQDRQHTLKHFEHVRMVDPKKAAQTRSQVWTHLRIYER 540

Qy 541 MNQSLSLYNNVPAVAEEIQDEVELLQKEQNYSDDVLANMISEPRISYGNDAIMPSTET 600
Db 541 MNQSLSLYNNVPAVAEEIQDEVELLQKEQNYSDDVLANMISEPRISYGNDAIMPSTET 600

Qy 601 KTTVELLPVNGEFLSDLOLPWHSFGADSPANTENEVEPVDARPAADRGLITTRPGSGLN 660
Db 601 KTTVELLPVNGEFLSDLOLPWHSFGADSPANTENEVEPVDARPAADRGLITTRPGSGLN 660

Qy 661 IKTEETSEVKMDAEFRHDSGYEVHHOKLVFFAEDVGSNGKAIIGLMVGWGVATVITVL 720
Db 661 IKTEETSEVKMDAEFRHDSGYEVHHOKLVFFAEDVGSNGKAIIGLMVGWGVATVITVL 720

Qy 721 VMLK$KQYTSIHGGVVEVDAATPEERHLSKMQQNGYENPTYKFFEQMQN 770

```

```

Db 721 VMLK$KQYTSIHGGVVEVDAATPEERHLSKMQQNGYENPTYKFFEQMQN 770

RESULT 2
AAR41546
ID AAR41546 standard; Protein; 770 AA.
XX AAR41546;
XX AC AAR41546;
XX DT 15-MAR-1994 (first entry)
XX DE Mutated APP770 exon 17 protein fragment.
XX KW Probe; mutation; exon 17; amyloid precursor protein; APP770;
XX KW substitution; progressive presenile dementia; Alzheimer's disease;
XX KW cerebral haemorrhage; cerebral amyloid angiopathy; 4 kD fragment;
XX KW blood vessels; brain parenchyma; assay; processing; plaque.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT Misc-difference 692
FT /note= "Position of mutation which causes abnormal
FT processing of APP770"
XX PN EP561087-A.
XX PD 22-SEP-1993.
XX PF 20-MAR-1992; 92EP-0400771.
XX PR 20-MAR-1992; 92EP-0400771.
XX PA (INNO-) INNOGENETICS NV SA.
XX PI Cras P, Hendriks L, Martin J, Van Broeckhoven C;
XX DR WPI; 1993-296442/38.
XX DR N-PSDB; AAQ48860.
XX PT New mutant form of beta-amyloid polypeptide - related to
XX PT development of cerebral haemorrhage and Alzheimer's disease, also
XX PT corresp. nucleic acid, vectors, host cells and antibodies
XX PS Disclosure; Fig 4; 21pp; English.
XX CC This sequence is encoded by exon 17 of the amyloid precursor protein
XX CC APP770. A mutation at position 2075 of the DNA sequence, a C>G
XX CC substitution, causes the substitution of Ala for Gly at codon 692.
XX CC This mutation has been detected in related patients with progressive
XX CC presenile dementia (Alzheimer's disease) or cerebral haemorrhage due
XX CC to cerebral amyloid angiopathy. It may be responsible for the
XX CC deposition of a 4 kD proteolytic fragment of APP in blood vessel
XX CC walls and brain parenchyma. Probes specific for the mutation (see
XX CC also AAQ48858-59) can be used to assay mRNA encoding substances which
XX CC cause abnormal processing of APP related to plaque formation, and to
XX CC detect this specific mutation.
XX SQ Sequence 770 AA;

Query Match 100.0%; Score 4058; DB 14; Length 770;
Best Local Similarity 100.0%; Pred. No. 5.3e-289;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPOIAMECGRLLNMHMNVONGKWDSPSGTK 60
Db 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPOIAMECGRLLNMHMNVONGKWDSPSGTK 60

Qy 61 TCIDTKEGILQYCEVYPELQITNVVEANQPVTIQNCKRGRKCKTHPHFVPIYRCLVG 120
Db 61 TCIDTKEGILQYCEVYPELQITNVVEANQPVTIQNCKRGRKCKTHPHFVPIYRCLVG 120

```

```

QY 121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSKSTNLHDYGMLLPCGIDKFR 180
DB 121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSKSTNLHDYGMLLPCGIDKFR 180
QY 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEEVAEVEE 240
DB 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEEVAEVEE 240
QY 241 EADDEDEDEGDEVEEAEPYEATERTTSIATTTTTTSTESVEEVVREVCSQAETGPC 300
DB 241 EADDEDEDEGDEVEEAEPYEATERTTSIATTTTTTSTESVEEVVREVCSQAETGPC 300
QY 301 RAMISRWFVDTGKCAFFYGGCGGNRNFDEEYCMVCGSAMQSLLKTQEPPLARD 360
DB 301 RAMISRWFVDTGKCAFFYGGCGGNRNFDEEYCMVCGSAMQSLLKTQEPPLARD 360
QY 361 PVKLPPTAASTPDVADKYLETGDENEHAHFQKAKERLEAKHRMSQVMREWEAERQA 420
DB 361 PVKLPPTAASTPDVADKYLETGDENEHAHFQKAKERLEAKHRMSQVMREWEAERQA 420
QY 421 KNLPRKADKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
DB 421 KNLPRKADKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
QY 481 QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 540
DB 481 QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 540
QY 541 MNQSLSLLYNVPVAEEIQDEVELLQKEQYSDVLANMISEPRISYGNDAIMPSTLET 600
DB 541 MNQSLSLLYNVPVAEEIQDEVELLQKEQYSDVLANMISEPRISYGNDAIMPSTLET 600
QY 601 KTTVELLPVNGEFSLLDLPQWHSFGADSVPAANTEVEPEVDARPAADRGLTTRPGSGLTN 660
DB 601 KTTVELLPVNGEFSLLDLPQWHSFGADSVPAANTEVEPEVDARPAADRGLTTRPGSGLTN 660
QY 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNGKGAIIGLMVGWVIATVIVITL 720
DB 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNGKGAIIGLMVGWVIATVIVITL 720
QY 721 VMLKKKQYTSIIHGVEVDAAVTPPEERHLSKMQONGYENPTYKFFQMOMN 770
DB 721 VMLKKKQYTSIIHGVEVDAAVTPPEERHLSKMQONGYENPTYKFFQMOMN 770

RESULT 3
AAR63442
ID AAR63442 standard; protein; 770 AA.
AC AC
XX XX
XX XX
DT 30-JUN-1995 (first entry)
DE Amyloid protein precursor APP.
XX XX
KW Amyloid protein precursor; APP; beta-amyloid protein; gelatinase A;
KW progelatinase A; decomposing agent; prophylaxis; Alzheimer's disease.
XX XX
OS Homo sapiens.
XX XX
FH Key
FT Region
FT 672..711
FT /note= "beta-AP 1-40 gelatinase A
FT cleavage sites"
FT
FT Cleavage-site 687..689
FT /label= gelatinase A
FT
FT Cleavage-site 701..702
FT /label= gelatinase A
FT
FT Cleavage-site 705..706
FT /label= gelatinase A
XX XX
PN EP622079-A.
XX XX

```

```

PD 02-NOV-1994..
XX
PF 25-APR-1994; 94EP-0302924.
XX
PR 27-APR-1993; 93JP-0122207.
PR 25-FEB-1994; 94JP-0051133.
XX
PA (ORIY ) ORIENTAL YEAST CO LTD.
XX
PI Miyazaki K;
XX
WI; 1994-334379/42.
XX
Gelatinase A-contg. amyloid beta protein decomposing agent -
useful for prophylaxis or treatment of Alzheimer's disease
XX
Disclosure: Fig 1; 12pp; English.
XX
AAR63442 describes the amino acid sequence of the amyloid protein
precursor (APP), from which beta amyloid protein is derived (APB).
An APB decomposing agent having either gelatinase A, a limited
decomposate of gelatinase A, or progelatinase A as an active
ingredient was developed. This agent can be used in medicine for
prophylaxis and for the treatment of Alzheimer's disease.
XX
SQ Sequence 770 AA;
Query Match 100.0%; Score 4058; DB 15; Length 770;
Best Local Similarity 100.0%; Pred. No. 5.3e-289;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPOIAMFCGRLNMHMNVONGKWDSPSGTK 60
DB 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPOIAMFCGRLNMHMNVONGKWDSPSGTK 60
QY 61 TCIDTKEGILQYCEVPELQITNVVEANQPVTIQNMCKRGKQCKTHPHFVIPYRCVLG 120
DB 61 TCIDTKEGILQYCEVPELQITNVVEANQPVTIQNMCKRGKQCKTHPHFVIPYRCVLG 120
QY 121 EFVSALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSKSTNLHDYGMLLPCGIDKFR 180
DB 121 EFVSALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSKSTNLHDYGMLLPCGIDKFR 180
QY 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEEVAEVEE 240
DB 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEEVAEVEE 240
QY 241 EADDEDEDEGDEVEEAEPYEATERTTSIATTTTTTSTESVEEVVREVCSQAETGPC 300
DB 241 EADDEDEDEGDEVEEAEPYEATERTTSIATTTTTTSTESVEEVVREVCSQAETGPC 300
QY 301 RAMISRWFVDTGKCAFFYGGCGGNRNFDEEYCMVCGSAMQSLLKTQEPPLARD 360
DB 301 RAMISRWFVDTGKCAFFYGGCGGNRNFDEEYCMVCGSAMQSLLKTQEPPLARD 360
QY 361 PVKLPPTAASTPDVADKYLETGDENEHAHFQKAKERLEAKHRMSQVMREWEAERQA 420
DB 361 PVKLPPTAASTPDVADKYLETGDENEHAHFQKAKERLEAKHRMSQVMREWEAERQA 420
QY 421 KNLPRKADKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
DB 421 KNLPRKADKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
QY 481 QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 540
DB 481 QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 540
QY 541 MNQSLSLLYNVPVAEEIQDEVELLQKEQYSDVLANMISEPRISYGNDAIMPSTLET 600
DB 541 MNQSLSLLYNVPVAEEIQDEVELLQKEQYSDVLANMISEPRISYGNDAIMPSTLET 600
QY 601 KTTVELLPVNGEFSLLDLPQWHSFGADSVPAANTEVEPEVDARPAADRGLTTRPGSGLTN 660
DB 601 KTTVELLPVNGEFSLLDLPQWHSFGADSVPAANTEVEPEVDARPAADRGLTTRPGSGLTN 660

```

Db 601 KTTVELLPVNGEESLDDLPWHSFGADSVDPANTENEVEPVDARPAADRLTTRPGSLTN 660
QY 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVIATVITVL 720
Db 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVIATVITL 720
QY 721 VMLKKQYTSIHGGVVEVDAAVTPPEERHLSKMQQNGYENPTYKFFEQMQN 770
Db 721 VMLKKQYTSIHGGVVEVDAAVTPPEERHLSKMQQNGYENPTYKFFEQMQN 770

RESULT 4
AAW40130 standard; Protein: 770 AA.
ID AAW40130;
AC AAW40130;
XX
DT 03-JUN-1998 (first entry)
DE Human APP770 protein.
XX
KW Amyloid-beta peptide; A-beta; beta-AP; senile plaque; angiopathy;
KW brain; membrane-spanning glycoprotein; beta-amyloid precursor protein;
KW APP770; chromosome 21; human; Alzheimers disease; AD; amyloid filament;
KW treatment; disease; Down's syndrome; hereditary cerebral haemorrhage.
XX
OS Homo sapiens.
XX
PN WO9748983-AL.
XX
PD 24-DEC-1997.
XX
PF 18-JUN-1997; 97WO-US10601.
XX
PR 18-JUN-1996; 96US-0665649.
XX
PA (ATHE-) ATHENA NEUROSCIENCES INC.
PA (BGHM) BRIGHAM & WOMENS HOSPITAL.
XX
PI Citron M, Schenk D, Selkoe DJ, Seubert PA;
XX
DR WPI: 1998-063287/06.
DR N-PSDB; AAV10322.
XX
PT Identifying compounds that alter cellular production of amyloid-beta
PT 42 fragment - in vitro or in transgenic animal models, potentially
PT useful for treatment of Alzheimer's and other amyloid deposition
PT diseases
XX
PS Disclosure; Fig 10; 86pp; English.
XX
CC This sequence represents the human beta-amyloid precursor protein
CC APP770, which is a membrane-spanning glycoprotein encoded by a gene
CC on the long arm of chromosome 21. A fragment of the APP protein is known
CC as the amyloid-beta peptide (A-beta), also known as the beta-AP peptide,
CC which forms the subunit of the amyloid filaments comprising senile
CC (amyloid) plaques and the amyloid deposits in small cerebral and
CC meningeal blood vessels (amyloid angiopathy). The A-beta peptide can be
CC a 39-43 amino acid fragment. This invention provides methods of screening
CC compounds for their ability to alter the production of the A-beta
CC peptide, which is composed of >41 amino acids, alone, or in combination
CC with the A-beta peptide composed of 40 amino acids or less. Such agents
CC that reduce the production of the A-beta peptide are potentially useful
CC for treatment of Alzheimers Disease or other diseases involving amyloid
CC deposition such as Down's syndrome, hereditary cerebral haemorrhage with
CC amyloidosis of Dutch type and advanced aging of the brain.
XX
SQ Sequence 770 AA;

Query Match 100.0%; Score 4058; DB 19; Length 770;
Best Local Similarity 100.0%; Pred. No. 5,3e-289;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLPGLALLLLAANTARALEVPTDGNAGLLABPQIAMFCGRLLNMHNVQNGKWDSDPSGTK 60
Db 1 MLPGLALLLLAANTARALEVPTDGNAGLLABPQIAMFCGRLLNMHNVQNGKWDSDPSGTK 60
QY 61 TCIDTKEGILQYCOEVPPELOITNVVEANQVPTIQNCKRGRKCKTHPHFVPIPYRCLVG 120
Db 61 TCIDTKEGILQYCOEVPPELOITNVVEANQVPTIQNCKRGRKCKTHPHFVPIPYRCLVG 120
QY 121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPGCIDKFR 180
Db 121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPGCIDKFR 180
QY 181 GVEFVCCPLABESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEAEVEEVEE 240
Db 181 GVEFVCCPLABESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEAEVEEVEE 240
QY 241 EADDDDEDDGDEVEEAEPEEYEEATERTTSTATTTTTTTSTESVEEVREVCSQAETGPC 300
Db 241 EADDDDEDDGDEVEEAEPEEYEEATERTTSTATTTTTTTSTESVEEVREVCSQAETGPC 300
QY 301 RAMISRWFYDTEGKCAPFFYGGCGGNRNFDTEYCMVCGSAMSQSLKTTQEPPLARD 360
Db 301 RAMISRWFYDTEGKCAPFFYGGCGGNRNFDTEYCMVCGSAMSQSLKTTQEPPLARD 360
QY 361 PVKLPPTTAASTPDVADKYLETTPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAEARQA 420
Db 361 PVKLPPTTAASTPDVADKYLETTPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAEARQA 420
QY 421 KNLPKADKKAVIQHFOEKVESLEQEAANERQQLVETHMARVEAMLNDRRRLALENYITAL 480
Db 421 KNLPKADKKAVIQHFOEKVESLEQEAANERQQLVETHMARVEAMLNDRRRLALENYITAL 480
QY 481 QAVPPRPRHVENMLKYYRAEQKDRQHTLKHFHVRVMDPKAAQIRSQVMTLRLVIER 540
Db 481 QAVPPRPRHVENMLKYYRAEQKDRQHTLKHFHVRVMDPKAAQIRSQVMTLRLVIER 540
QY 541 MNQSLSLLYNPVPAVEETQDEVDLQKEQYSDVLANMISEPRISYNDALMPSLTET 600
Db 541 MNQSLSLLYNPVPAVEETQDEVDLQKEQYSDVLANMISEPRISYNDALMPSLTET 600
QY 601 KTTVELLPVNGEESLDDLPWHSFGADSVDPANTENEVEPVDARPAADRLTTRPGSLTN 660
Db 601 KTTVELLPVNGEESLDDLPWHSFGADSVDPANTENEVEPVDARPAADRLTTRPGSLTN 660
QY 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVIATVITVL 720
Db 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVIATVITL 720
QY 721 VMLKKQYTSIHGGVVEVDAAVTPPEERHLSKMQQNGYENPTYKFFEQMQN 770
Db 721 VMLKKQYTSIHGGVVEVDAAVTPPEERHLSKMQQNGYENPTYKFFEQMQN 770

RESULT 5
AAW97996 standard; Protein: 770 AA.
ID AAW97996;
AC AAW97996;
XX
DT 21-JUN-1999 (first entry)
DE Human amyloid precursor protein.
XX
KW Amyloid precursor protein; App; human; gene targeting;
KW homologous recombination; transgenic mouse; transgenic animal;
KW animal model; Alzheimer's disease.
XX
OS Homo sapiens.
XX
FH Key 672..711
FT Domain /note="beta-amyloid domain"
XX

PN WO9909150-A1.
 XX 25-FEB-1999.
 XX 18-AUG-1997; 97WO-US14507.
 XX 18-AUG-1997; 97WO-US14507.
 XX (FARB) BAYER CORP.
 XX Wirak DO;
 XX WPI; 1999-181029/15.
 DR Modification of target nucleic acids - by homologous recombination,
 XX particularly for introducing a humanised amyloid precursor
 XX protein gene into rodents for producing models of Alzheimer's
 XX disease
 XX Disclosure; Page 85-88; 209pp; English.
 XX This polypeptide comprises human amyloid precursor protein (hAPP).
 XX The invention provides a novel gene targeting strategy that
 XX facilitates the introduction of one or more specific mutations
 XX into any gene in a single double reciprocal homologous recombination
 XX step. The method has been used particularly for introducing a
 XX humanised APP gene into rodents for producing animal models of
 XX Alzheimer's disease (AD). 4 Independent lines of transgenic mice
 XX (lines ES5007, ES5103, ES5401 and ES5403) have been created using
 XX the novel gene targeting technique applied to embryonic stem cells.
 XX In each line, the mouse APP gene was modified to encode a
 XX mouse/human hybrid (m/hAPP) where amino acid residues 666-770 of
 XX APP770 are encoded by human cDNA sequences instead of mouse genomic
 XX exons (exons 16-18). Within these residues, only 3 amino acid
 XX differences exist between the mouse and human proteins, i.e.
 XX Gly-676 to Arg, Phe-681 to Thr and Arg-684 to His. This exon-cDNA
 XX fusion gene therefore encodes an APP containing a humanised
 XX beta-amyloid domain. Swedish, London, Swedish/London and stop
 XX mutations have also been introduced. Targetting vector sequences
 XX are provided (see AAX24730-33).
 XX Sequence 770 AA;
 SQ

Query Match 100.0%; Score 4058; DB 20; Length 770;
 Best Local Similarity 100.0%; Pred. No. 5.3e-289;
 Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLPGLALLLLAATARALEVPTDGNAGLLAEPQIAMFCGLRNHNMVQNGKWDSPSGTK 60
 DB 1 MLPGLALLLLAATARALEVPTDGNAGLLAEPQIAMFCGLRNHNMVQNGKWDSPSGTK 60
 QY 61 TCIDTREGILQYQEVYPELQITNVVEANOPVTIQNCKRGRKCKTHPEVTPYRCLV 120
 DB 61 TCIDTREGILQYQEVYPELQITNVVEANOPVTIQNCKRGRKCKTHPEVTPYRCLV 120
 QY 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHVTAKETCSKSTNLHDYGMLLPCGIDKFR 180
 DB 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHVTAKETCSKSTNLHDYGMLLPCGIDKFR 180
 QY 181 GVFEVCCPLAEESDNVDSADAEDSDVVMWGGADTDYADGSEDKVVEAEVEAEVEE 240
 DB 181 GVFEVCCPLAEESDNVDSADAEDSDVVMWGGADTDYADGSEDKVVEAEVEAEVEE 240
 QY 241 EADDEDEDEGDEVEEAEPYEATERTSIATTTTTTSTVESVEEVVREVCSEQATGFC 300
 DB 241 EADDEDEDEGDEVEEAEPYEATERTSIATTTTTTSTVESVEEVVREVCSEQATGFC 300
 QY 301 RAMISRWFVDVTEGKCAPFFYGGCGGNRNFDTEECMAVCGSAMQSLLKTTQEP LARD 360
 DB 301 RAMISRWFVDVTEGKCAPFFYGGCGGNRNFDTEECMAVCGSAMQSLLKTTQEP LARD 360
 QY 361 PVKLPPTAASTPDVADKYLETPGDENEHAHFQKAKERLEAKHRMSQVMRWEAEARQA 420
 DB 361 PVKLPPTAASTPDVADKYLETPGDENEHAHFQKAKERLEAKHRMSQVMRWEAEARQA 420

Db 361 PVKLPPTAASTPDVADKYLETPGDENEHAHFQKAKERLEAKHRMSQVMRWEAEARQA 420
 QY 421 KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENY ITAL 480
 Db 421 KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENY ITAL 480
 QY 481 QAVPPRPRIHVFNMLKKYVRAEQDKRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIER 540
 Db 481 QAVPPRPRIHVFNMLKKYVRAEQDKRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIER 540
 QY 541 MNQSLSLYNVPAVAEEIODEVDELQKQNYSDVLANNISEPRIISYNDALMPSLTET 600
 Db 541 MNQSLSLYNVPAVAEEIODEVDELQKQNYSDVLANNISEPRIISYNDALMPSLTET 600
 QY 601 KTTVELLPVNGEFSDDLQPHWSFGADSVPAANTENEVEPVDARPAADRLGTTTPGSGLTN 660
 Db 601 KTTVELLPVNGEFSDDLQPHWSFGADSVPAANTENEVEPVDARPAADRLGTTTPGSGLTN 660
 QY 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEEDVGSNKGAIIGLMVGGVVATVITVL 720
 Db 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEEDVGSNKGAIIGLMVGGVVATVITVL 720
 QY 721 VMLKKQYTSIHGGVVEVDAAYTPPERHLSKMQQNGYENPTYKFFEQMQN 770
 Db 721 VMLKKQYTSIHGGVVEVDAAYTPPERHLSKMQQNGYENPTYKFFEQMQN 770
 RESULT 6
 AAE10648
 ID AAE10648 standard; Protein; 770 AA.
 XX AAE10648;
 AC AAE10648;
 XX 10-DEC-2001 (first entry)
 DT Human amyloid protein precursor 770 (APP770) isoform.
 DE Human; aspartyl protease 1; Aspl; amyloid precursor protein 770; APP770;
 KW Alzheimer's disease; AD; dementia; neurofibrillary tangle; gliosis;
 KW amyloid plaque; neuronal loss; proteolytic; neurotropic; neuroprotective.
 XX Homo sapiens.
 OS GB2357767-A.
 PN 04-JUL-2001.
 XX 22-SEP-2000; 2000GB-0023315.
 PF 23-SEP-1999; 99US-0155493.
 PR 23-SEP-1999; 99US-0404133.
 PR 23-SEP-1999; 99WO-US20881.
 PR 13-OCT-1999; 99US-0416901.
 PR 06-DEC-1999; 99US-0169232.
 XX (PHAA) PHARMACIA & UPJOHN CO.
 PA Bienkowski MJ, Gurney M;
 PI WPI; 2001-444208/48.
 DR N-PSDB; AAD17897.
 XX Polypeptide comprising fragments of human aspartyl protease with
 PT amyloid precursor protein processing activity and alpha-secretase
 PT activity, for identifying modulators useful in treating Alzheimer's
 PT disease -
 XX Example 8; Page 142-144; 187pp; English.
 PS The patent discloses human aspartyl protease 1 (hu-Aspl) or modified
 XX Aspl proteins which lack transmembrane domain or amino terminal
 CC domain or cytoplasmic domain and retains alpha-secretase activity
 CC and amyloid protein precursor (APP) processing activity. The proteins

CC of the invention are useful for assaying hu-Asp1 alpha-secretase
CC activity, which in turn is useful for identifying modulators of
CC hu-Asp1 alpha-secretase activity, where modulators that increase
CC hu-Asp1 alpha-secretase activity are useful for treating Alzheimer's
CC disease (AD) which causes progressive dementia with consequent
CC formation of amyloid plaques, neurofibrillary tangles, gliosis and
CC neuronal loss. Hu-Asp1 protease substrate is useful for assaying
CC hu-Asp1 proteolytic activity, by contacting hu-Asp1 protein with
CC the substrate under acidic conditions and determining the level of
CC hu-Asp1 proteolytic activity. The present sequence is amyloid protein
CC precursor 770 (APP770) isoform from human.

XX SQ Sequence 770 AA;

Query Match 100.0%; Score 4058; DB 22; Length 770;
Best Local Similarity 100.0%; Pred. No. 5.3e-289;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MLPGLALLLLAATARALEVPTDGNAGLLAEPQIAMEFCGRLNHMHYQNGKWDSPSGTK 60
Db 1 MLPGLALLLLAATARALEVPTDGNAGLLAEPQIAMEFCGRLNHMHYQNGKWDSPSGTK 60
Qy 61 TCIDTKEGILQYCEVYPELQITNVVEANQPVTTIQNCKGRKCKTHPHFVPIYRCLVG 120
Db 61 TCIDTKEGILQYCEVYPELQITNVVEANQPVTTIQNCKGRKCKTHPHFVPIYRCLVG 120
Qy 121 EFVSDALLVPDKCFKHQERMDVCETHLHHTVAKETCEKSTNLHDYGMLLPGCIDKFR 180
Db 121 EFVSDALLVPDKCFKHQERMDVCETHLHHTVAKETCEKSTNLHDYGMLLPGCIDKFR 180
Qy 181 GVEFYCCPLAESDNVDSADAEEDSDVWGGADTDYADGSEDKVVEAEVEEVEE 240
Db 181 GVEFYCCPLAESDNVDSADAEEDSDVWGGADTDYADGSEDKVVEAEVEEVEE 240
Qy 241 EADDEDEDGDEVEEAEPEYBEATERTTSIATTTTTTSTESVEEVREVCSEAEAGPC 300
Db 241 EADDEDEDGDEVEEAEPEYBEATERTTSIATTTTTTSTESVEEVREVCSEAEAGPC 300
Qy 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNFTDEYCMVCGSAMSOSLLKTTQEP LARD 360
Db 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNFTDEYCMVCGSAMSOSLLKTTQEP LARD 360
Qy 361 PVKLTPTAASPDVADKYLETTPGDENEHAHFQKAKERLEAKHRMSQVMREWEAEERQA 420
Db 361 PVKLTPTAASPDVADKYLETTPGDENEHAHFQKAKERLEAKHRMSQVMREWEAEERQA 420
Qy 421 KNLPRADKKAVIQHFQKVESLEQEAANERQOLVETHMARVEAMLNDRRLALENYITL 480
Db 421 KNLPRADKKAVIQHFQKVESLEQEAANERQOLVETHMARVEAMLNDRRLALENYITL 480
Qy 481 QAVPPRPRHVNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKAAQIRSQVTHLRVIER 540
Db 481 QAVPPRPRHVNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKAAQIRSQVTHLRVIER 540
Qy 541 MNQSLSLLYNPAVAEEIQDEVELLOKEQYSDVLANMISEPRISYGNDAIMPSTTET 600
Db 541 MNQSLSLLYNPAVAEEIQDEVELLOKEQYSDVLANMISEPRISYGNDAIMPSTTET 600
Qy 601 KTTVELLPVNGEFLDDLPWHSFGADSVPAANTEVEPVDARPAADGLTRPGSGLTN 660
Db 601 KTTVELLPVNGEFLDDLPWHSFGADSVPAANTEVEPVDARPAADGLTRPGSGLTN 660
Qy 661 IKTEISEVKMDAEPHSDGSEVHHQKLVFFAEDVGSNKGAIIGLVGGVVIATVITL 720
Db 661 IKTEISEVKMDAEPHSDGSEVHHQKLVFFAEDVGSNKGAIIGLVGGVVIATVITL 720
Qy 721 VMLKKQKQTSIHGGVVEVDAVTPPERHLSKMQQNGYENPTYKFFEQMN 770
Db 721 VMLKKQKQTSIHGGVVEVDAVTPPERHLSKMQQNGYENPTYKFFEQMN 770

RESULT 7
AAE11762

ID AAE11762 standard; Protein; 770 AA.
XX AAE11762;
AC AAE11762;
XX 18-DEC-2001 (first entry)
DT Human amyloid precursor protein (APP).
XX Human; neuroprotective; nootropic; immunostimulant; Alzheimer's disease;
KW anticonvulsant; vaccine; gene therapy; Pick's disease; antidiabetic;
KW systemic amyloidosis; maturity onset diabetes; Parkinson's disease;
KW Huntington's disease; fronto-temporal dementia; encephalopathy; ALS;
KW amyotrophic lateral sclerosis; amyloid precursor protein; APP.
XX Homo sapiens.
OS Homo sapiens.
XX Key Location/Qualifiers
FH Peptide 1..18
FT /label= Signal_peptide
FT Protein 19..770
FT /label= Mature_human_AAP_protein
FT Domain 18..700
FT /label= Extracellular_domain
FT Region 672..714
FT /note= "Abeta-42/43 core peptide"
FT Domain 700..723
FT /label= Transmembrane_domain
FT Region 714..770
FT /note= "C-100 fragment"
FT Domain 723..770
FT /label= Intracellular_domain
XX WO200162284-A2.
PN 30-AUG-2001.
PD 19-FEB-2001; 2001WO-DK00113.
XX 21-FEB-2000; 2000DK-0000265.
PR 01-MAR-2000; 2000US-186295P.
XX (WEBI-) M & E BIOTECH AS.
PA Birk P, Jensen MR, Nielsen KG;
PI WPI; 2001-589796/66.
XX N-PSDB; AAD18754.
DR In vivo down-regulation of amyloid protein for the treatment of
XX Alzheimer's, comprises presenting an amyloidogenic polypeptide or its
PT subsequence and/or at least one analogue of the amyloidogenic
PT polypeptide to the immune system -
XX Claim 23; Page 113-116; 120pp; English.
PS The invention relates to a method for in vivo down-regulation of amyloid
XX protein such as beta amyloid (Abeta) in an animal, including human. The
XX method comprising presenting to the animal's immune system an
XX immunogenically effective amount of at least one amyloidogenic protein
XX or its subsequence and/or at least one analogue of the amyloidogenic
XX polypeptide. The amyloidogenic protein or its subsequence, and its
XX analogue is useful for the preparation of an immunogenic composition
XX comprising an adjuvant for down-regulating amyloid in an animal. They are
XX also useful in the treatment, prophylaxis or amelioration of Alzheimer's
XX disease or other diseases characterised by amyloid deposits. They are
XX also useful in the treatment of systemic amyloidosis, maturity onset
XX diabetes, Parkinson's disease, Huntington's disease, fronto-temporal
XX dementia, amyotrophic lateral sclerosis (ALS), Pick's disease and
XX prion-related transmissible spongiform encephalopathies. They are also
XX useful for inducing production of antibodies against an amyloidogenic
XX polypeptide. The present sequence is human amyloid precursor protein
XX (APP).

SQ Sequence 770 AA;

Query Match 100.0%; Score 4058; DB 22; Length 770;
 Best Local Similarity 100.0%; Pred. No. 5.3e-289;
 Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLPGLALLLAATAARALEVPTDGNAGLLAEPOIAMFCGRLLNMHNVQNGKWDSDPSGK 60
 Db 1 MLPGLALLLAATAARALEVPTDGNAGLLAEPOIAMFCGRLLNMHNVQNGKWDSDPSGK 60

Qy 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTIONCKRGRKCKTHPHFVPIYRCLVG 120
 Db 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTIONCKRGRKCKTHPHFVPIYRCLVG 120

Qy 121 EFVSDALLVPDKCKFLHOERMDVCETHLHWHVTAKETCSEKSTNLHDYGMLLPCGIDKFR 180
 Db 121 EFVSDALLVPDKCKFLHOERMDVCETHLHWHVTAKETCSEKSTNLHDYGMLLPCGIDKFR 180

Qy 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEAEVEEVAEVEE 240
 Db 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEAEVEEVAEVEE 240

Qy 241 EADDEDEDEGDEVEEAEPEYEATERTTSIATTTTTTESVEEVVREVCSEAETGPC 300
 Db 241 EADDEDEDEGDEVEEAEPEYEATERTTSIATTTTTTESVEEVVREVCSEAETGPC 300

Qy 301 RAMLSRWYFDVTGKCAPFYGGCGGNRNFDTEECYMAVCGSAMQSLLKTTQEP LARD 360
 Db 301 RAMLSRWYFDVTGKCAPFYGGCGGNRNFDTEECYMAVCGSAMQSLLKTTQEP LARD 360

Qy 361 PVKLPTTAASTPDVADKYLETGPDENEHAFQKAKERLEAKHRMSQVMEWEAEAOA 420
 Db 361 PVKLPTTAASTPDVADKYLETGPDENEHAFQKAKERLEAKHRMSQVMEWEAEAOA 420

Qy 421 KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
 Db 421 KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480

Qy 481 QAYVPRPRHVNLMKKYVRAEQKDRQHTLKHFEHVRMVDPKAAQIRSOVMTHLRVIYER 540
 Db 481 QAYVPRPRHVNLMKKYVRAEQKDRQHTLKHFEHVRMVDPKAAQIRSOVMTHLRVIYER 540

Qy 541 MNQSLLYNVPAAVEIQDEVDLQKQNYSDVLANMISEPRISYNDALMPSLTET 600
 Db 541 MNQSLLYNVPAAVEIQDEVDLQKQNYSDVLANMISEPRISYNDALMPSLTET 600

Qy 601 KTTVELLPVNGEFLDLPWHSFGADSVPAANTENEVEPVDARPAADRLGTTTPGSGLTN 660
 Db 601 KTTVELLPVNGEFLDLPWHSFGADSVPAANTENEVEPVDARPAADRLGTTTPGSGLTN 660

Qy 661 IKTEETSEVKMDAEFRHDSGYEVHHOKLVFFAEADVGSNKGAIIGLMVGWVIATVITL 720
 Db 661 IKTEETSEVKMDAEFRHDSGYEVHHOKLVFFAEADVGSNKGAIIGLMVGWVIATVITL 720

Qy 721 VMLKKQYTSIHGGVVEVDAVTPPERHLSKMQQNGYENPTYKFFEQMON 770
 Db 721 VMLKKQYTSIHGGVVEVDAVTPPERHLSKMQQNGYENPTYKFFEQMON 770

RESULT 8

AAE06893

ID AAE06893 standard; Protein: 770 AA.

XX AC AAE06893;

XX DT

XX 23-OCT-2001 (first entry)

DE Human amyloid precursor protein 770 (APP770) isoform.

XX

KW Human; aspartyl protease; beta-amyloid precursor protein 770; APP770;
 KW beta-secretase; Alzheimer's disease; dementia; amyloid plaque; gliosis;
 KW neurofibrillary tangle; neuronal loss; amyloid-beta peptide; neurotropic;
 KW neuroprotective; antisense therapy; gene therapy; chromosome 21.

XX

OS Homo sapiens.

XX WO200150829-A2.

XX 19-JUL-2001.

XX 09-MAY-2001; 2001WO-IB00799.

XX 09-MAY-2001; 2001WO-IB00799.

XX (BIEN/) BIENKOWSKI M J.

XX (GURN/) GURNEY M E.

XX (HEIN/) HEINRIKSON R L.

XX (PARO/) PARODI L A.

XX (YANR/) YAN R.

XX

PI Bienkowski MJ, Gurney ME, Heinrikson RL, Parodi LA, Yan R;

XX WPI; 2001-483072/52.

XX DR N-PSDB; AAD13278.

XX

PT Novel purified polypeptide comprising fragment of mammalian aspartyl

PT protease 2, lacking Asp2 transmembrane domain and retaining beta

PT secretase activity of Asp2 useful for identifying inhibitors of Asp2

PT activity -

PS Claim 8; Page 171-173; 185pp; English.

XX

CC The invention relates to human aspartyl proteases (Hu-Asp), beta-amyloid

CC precursor protein (APP) isoforms and their corresponding DNA molecules.

CC Human aspartyl proteases can act as beta-secretase proteases useful for

CC treating Alzheimer's disease. APP isoforms are useful for identifying

CC modulators of amyloid-beta peptide production, for use in designing

CC therapeutics for the treatment and prevention of Alzheimer's disease,

CC dementia, formation of amyloid plaques, neurofibrillary tangles, gliosis

CC and neuronal loss. APP isoforms are also used in methods for identifying

CC inhibitors and modulators of human Asp2 activity. The invention relates

CC to a method for identifying agents that modulate the activity of human

CC aspartyl protease Asp2. Amyloid-beta peptides obtained from APP are used

CC as a means to screen in cellular assays for the inhibitors of beta- and

CC gamma- secretase. Hu-Asp DNA fragments are useful as probes or primers in

CC polymerase chain reactions (PCR). The probes are useful for detecting

CC Hu-Asp nucleic acids in in vitro assays and in Northern and Southern

CC blots. The present sequence is human wild-type amyloid precursor

CC protein 770 (APP770) isoform. APP770 gene is localised of chromosome 21.

XX SQ Sequence 770 AA;

Query Match 100.0%; Score 4058; DB 22; Length 770;

Best Local Similarity 100.0%; Pred. No. 5.3e-289;

Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLPGLALLLAATAARALEVPTDGNAGLLAEPOIAMFCGRLLNMHNVQNGKWDSDPSGK 60

Db 1 MLPGLALLLAATAARALEVPTDGNAGLLAEPOIAMFCGRLLNMHNVQNGKWDSDPSGK 60

Qy 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTIONCKRGRKCKTHPHFVPIYRCLVG 120

Db 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTIONCKRGRKCKTHPHFVPIYRCLVG 120

Qy 121 EFVSDALLVPDKCKFLHOERMDVCETHLHWHVTAKETCSEKSTNLHDYGMLLPCGIDKFR 180

Db 121 EFVSDALLVPDKCKFLHOERMDVCETHLHWHVTAKETCSEKSTNLHDYGMLLPCGIDKFR 180

Qy 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEAEVEEVAEVEE 240

Db 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEAEVEEVAEVEE 240

Qy 241 EADDEDEDEGDEVEEAEPEYEATERTTSIATTTTTTESVEEVVREVCSEAETGPC 300

Db 241 EADDEDEDEGDEVEEAEPEYEATERTTSIATTTTTTESVEEVVREVCSEAETGPC 300

QY 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNFTDEEYCMVCGSAMQSLLKTTQEP LARD 360
 DB 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNFTDEEYCMVCGSAMQSLLKTTQEP LARD 360
 QY 361 PVKLPPTAASTPDVADKYLETDPGDENHAHFQKAKERLEAKHRERMSQVMEWEAE AROA 420
 DB 361 PVKLPPTAASTPDVADKYLETDPGDENHAHFQKAKERLEAKHRERMSQVMEWEAE AROA 420
 QY 421 KNLPAKDKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALEN VITAL 480
 DB 421 KNLPAKDKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALEN VITAL 480
 QY 481 QAVPPRRHVFNM LKKYVRAEQDRQHTLKHFEHVRMVDPKKAAQIRSQVMT HRLVIYER 540
 DB 481 QAVPPRRHVFNM LKKYVRAEQDRQHTLKHFEHVRMVDPKKAAQIRSQVMT HRLVIYER 540
 QY 541 MNQSLSLYNNVPAVAEIQDEVDLQEQNYSDVYLANMISEPRISYGN DMLMPSLTET 600
 DB 541 MNQSLSLYNNVPAVAEIQDEVDLQEQNYSDVYLANMISEPRISYGN DMLMPSLTET 600
 QY 601 KTTVELLPVNGEFSLDLQPHWSEFADSVPAANTENEVEPVDARPAADRGLTTRPGSGLTN 660
 DB 601 KTTVELLPVNGEFSLDLQPHWSEFADSVPAANTENEVEPVDARPAADRGLTTRPGSGLTN 660
 QY 661 IKTEIESEVKMDAEFRHDSGYEVHHQKLVFFAEVDGSKNGKAIIGLMVGGVVIATVIVITL 720
 DB 661 IKTEIESEVKMDAEFRHDSGYEVHHQKLVFFAEVDGSKNGKAIIGLMVGGVVIATVIVITL 720
 QY 721 VMLKKKQYTSIHGQVVEVDAVTPPEERHLSKMQONGYENPTYKFFEQMQN 770
 DB 721 VMLKKKQYTSIHGQVVEVDAVTPPEERHLSKMQONGYENPTYKFFEQMQN 770

RESULT 9
 AAU06622
 ID AAU06622 standard; Protein; 770 AA.

XX AAU06622;

DT 24-OCT-2001 (first entry)

XX Human partial Amyloid precursor protein, APP770.

KW Human; Aspartyl protease; Asp2; beta-secretase; neurotropic; neuroprotective; amyloid protein precursor; APP; Alzheimer's disease; amyloid-beta; Abeta; APP770.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 358 /note= "Encoded by GCC"

XX WO200149098-A2.

XX 12-JUL-2001.

XX 09-MAY-2001; 2001WO-IB00798.

XX 09-MAY-2001; 2001WO-IB00798.

XX (BIEN/) BIENKOWSKI M J.

PA (GURN/) GURNEY M E.

PA (HEIN/) HEINRIKSON R L.

PA (PARO/) PARODI L A.

XX (YANR/) YAN R.

PI Bienkowski MJ, Gurney ME, Heinrikson RL, Parodi LA, Yan R;

XX WPI; 2001-502549/55.

DR N-PSDB; AAS11549.

XX Novel purified polypeptide comprising fragment of mammalian aspartyl

PT protease 2, lacking Asp2 transmembrane domain and retaining beta
 PT secretase activity of Asp2 useful for identifying inhibitors of Asp2
 activity

XX Disclosure; Page 171-173; 185pp; English.

XX The invention relates to a purified polypeptide comprising a fragment of
 CC mammalian aspartyl protease (Asp2) protein which lacks the Asp2
 CC transmembrane domain and the Asp2 protein, and where the polypeptide and
 CC the fragment retain the beta-secretase activity of the mammalian Asp2
 CC protein. The invention also details polynucleotides for the Asp
 CC proteins and vectors expressing them, and a polypeptide (isoform of
 CC amyloid protein precursor (APP)) comprising the amino acid sequence of an
 CC APP or its fragment containing an APP cleavage site recognizable by a
 CC mammalian beta-secretase, and further comprising two lysine residues at
 CC the carboxyl terminus of the amino acid sequence of the mammalian APP or
 CC APP fragment. Also included in the invention are methods of identifying
 CC modulators or inhibitors of Asp2. Modulators and inhibitors of Asp2 are
 CC useful for treating Alzheimer's disease. APP is useful in methods for
 CC identifying inhibitors or modulators of human Asp2 activity and
 CC amyloid-beta (Abeta) peptide production. APP is also useful in designing
 CC therapeutics for the treatment or prevention of Alzheimer's disease.
 CC APP comprising the APP-Sw-beta-secretase peptide sequence (NLDA), which
 CC is associated with increased levels of Abeta processing is useful in
 CC assays relating the Alzheimer's research. The expression vector is useful
 CC for recombinantly expressing APP. Nucleic acids that hybridize to
 CC Asp oligonucleotides are useful as probes or primers. The probes are
 CC useful for detecting Hu-Asp nucleic acids in in vitro assays and in
 CC Northern and Southern blots. The present sequence is human APP770.

XX Sequence 770 AA;

Query Match 100.0%; Score 4058; DB 22; Length 770;
 Best Local Similarity 100.0%; Pred. No. 5.3e-289;
 Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPIAMFCGRLNNHNVQNGKWDSPSGTK 60

DB 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPIAMFCGRLNNHNVQNGKWDSPSGTK 60

QY 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVYRCVLG 120

DB 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVYRCVLG 120

QY 121 EFVSDALLVPDKCKFLHQRMDVCEVTHLHWHTVAKETSEKSTNLHDYGMLLPCGIDKFR 180

DB 121 EFVSDALLVPDKCKFLHQRMDVCEVTHLHWHTVAKETSEKSTNLHDYGMLLPCGIDKFR 180

QY 181 GVEFVCCPLAESDNVDSDADEDDSDVWVGADTDYADGSEDKVVEVAEEVEEVEE 240

DB 181 GVEFVCCPLAESDNVDSDADEDDSDVWVGADTDYADGSEDKVVEVAEEVEEVEE 240

QY 241 EADDDDEDDGDEVEAEPEEATERTSTATTTTTTSTESVEEVVRCVSEQAETGPC 300

DB 241 EADDDDEDDGDEVEAEPEEATERTSTATTTTTTSTESVEEVVRCVSEQAETGPC 300

QY 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNFTDEEYCMVCGSAMQSLLKTTQEP LARD 360

DB 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNFTDEEYCMVCGSAMQSLLKTTQEP LARD 360

QY 361 PVKLPPTAASTPDVADKYLETDPGDENHAHFQKAKERLEAKHRERMSQVMEWEAE AROA 420

DB 361 PVKLPPTAASTPDVADKYLETDPGDENHAHFQKAKERLEAKHRERMSQVMEWEAE AROA 420

QY 421 KNLPAKDKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALEN VITAL 480

DB 421 KNLPAKDKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALEN VITAL 480

QY 481 QAVPPRRHVFNM LKKYVRAEQDRQHTLKHFEHVRMVDPKKAAQIRSQVMT HRLVIYER 540

DB 481 QAVPPRRHVFNM LKKYVRAEQDRQHTLKHFEHVRMVDPKKAAQIRSQVMT HRLVIYER 540

QY 541 MNQSLSLYNNVPAVAEIQDEVDLQEQNYSDVYLANMISEPRISYGN DMLMPSLTET 600


```

Db 541 MNQSLSLLYNPVAAEEIQDEVDLQKEQNYSDVLANMISEPRISYGNDAIMPSTET 600
QY 601 KTTVELLPVNGEFLDLDLPWHSFGADSVSPANTENEVEPVDARPAADRGLTTRPGSGLTN 660
Db 601 KTTVELLPVNGEFLDLDLPWHSFGADSVSPANTENEVEPVDARPAADRGLTTRPGSGLTN 660
QY 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEVDGSKNGCAIICGLMVGGVVIATVITL 720
Db 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEVDGSKNGCAIICGLMVGGVVIATVITL 720
QY 721 VMLKKKQYTSIHGGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN 770
Db 721 VMLKKKQYTSIHGGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN 770

RESULT 10
AAE02600
ID AAE02600 standard; Protein: 770 AA.
XX
AC AAE02600;
XX
DT 10-AUG-2001 (first entry)
XX
DE Human amyloid precursor protein 770 (APP 770) protein.
XX
KW Human; alpha-secretase; amyloid precursor protein; APP; therapy;
KW Alzheimer's disease; antialzheimer's.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 358 /note= "Encoded by GGC"
FT
XX
XX WO200123533-A2.
XX
XX 05-APR-2001.
XX
XX 22-SEP-2000; 2000WO-US26080.
XX
XX 23-SEP-1999; 99US-0155493.
XX 23-SEP-1999; 99WO-US20881.
XX 13-OCT-1999; 99US-0416901.
XX 06-DEC-1999; 99US-0169232.
XX
XX (PHAA ) PHARMACIA & UPJOHN CO.
XX
XX Gurney M, Bienkowski MJ;
XX
XX WPI; 2001-290516/30.
XX N-PSDB; AAD06770.
XX
XX Enzymes that cleave the alpha-secretase site of the amyloid precursor
XX protein, useful for the treatment of Alzheimer's disease -
XX
XX Example 8; Page 170-172; 189pp; English.
XX
XX The present invention relates to enzymes for cleaving the alpha-
XX secretase site of the amyloid precursor protein (APP) and methods of
XX identifying those enzymes. The methods may be used to identify enzymes
XX that may be used to cleave the alpha-secretase cleavage site of the APP
XX protein. The enzymes may be used to treat or modulate the progress of
XX Alzheimer's disease. The present sequence is human APP 770 protein.
XX
XX Sequence 770 AA;
XX
XX Query Match 100.0%; Score 4058; DB 22; Length 770;
XX Best Local Similarity 100.0%; Pred. No. 5.3e-289;
XX Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MLPGLALLLAWTARALEVPTDGNAGLLAEQIAMFCGRLNMHNMVQNGKWDSPSGTK 60
|||||

```

```

Db 1 MLPGLALLLAWTARALEVPTDGNAGLLAEQIAMFCGRLNMHNMVQNGKWDSPSGTK 60
QY 61 TCIDTREGILQYCOEYVPELQITNVVEANQPVTIQNWCKRGKCKOCTHPHFVLPYRCLVG 120
|||||
Db 61 TCIDTREGILQYCOEYVPELQITNVVEANQPVTIQNWCKRGKCKOCTHPHFVLPYRCLVG 120
QY 121 EFVSDALLVPDKCFHJQERMDVCEPHLHWHHTVAKETSEKSTNLHDYGMLLPCGIDKFR 180
|||||
Db 121 EFVSDALLVPDKCFHJQERMDVCEPHLHWHHTVAKETSEKSTNLHDYGMLLPCGIDKFR 180
QY 181 GVEFVCCPLAEESDNVDSADAEEDSDVMWGADTDYADGSEDKVVVEAAEEVEAEVEE 240
|||||
Db 181 GVEFVCCPLAEESDNVDSADAEEDSDVMWGADTDYADGSEDKVVVEAAEEVEAEVEE 240
QY 241 EADDDDEDDGDEVEEEAEPEYEATERTTSIATTTTTTSTESVEEVVRCVSEQAETGPC 300
|||||
Db 241 EADDDDEDDGDEVEEEAEPEYEATERTTSIATTTTTTSTESVEEVVRCVSEQAETGPC 300
QY 301 RAMISRWYFDVTGKCAPFFYGGCGGNRNFTTEYCMVCGSAMQSLLKTTQBPPLARD 360
|||||
Db 301 RAMISRWYFDVTGKCAPFFYGGCGGNRNFTTEYCMVCGSAMQSLLKTTQBPPLARD 360
QY 361 PVKLPPTAASTPDVADVKYLETPGDENEHAFQKAKERLEAKHRERMSQVMREWEAEQA 420
|||||
Db 361 PVKLPPTAASTPDVADVKYLETPGDENEHAFQKAKERLEAKHRERMSQVMREWEAEQA 420
QY 421 KNLKADKAVIQHFQEKVESLEQEAANERQOLVETHMARVEAMLNDRRLALENYITAL 480
|||||
Db 421 KNLKADKAVIQHFQEKVESLEQEAANERQOLVETHMARVEAMLNDRRLALENYITAL 480
QY 481 QAVPRPRHVFNMLKKYVRAEQKQKQHTLKHFEHVMVDPKKAAQIRSQVMTHLRVIER 540
|||||
Db 481 QAVPRPRHVFNMLKKYVRAEQKQKQHTLKHFEHVMVDPKKAAQIRSQVMTHLRVIER 540
QY 541 MNQSLSLLYNPVAAEEIQDEVDLQKEQNYSDVLANMISEPRISYGNDAIMPSTET 600
|||||
Db 541 MNQSLSLLYNPVAAEEIQDEVDLQKEQNYSDVLANMISEPRISYGNDAIMPSTET 600
QY 601 KTTVELLPVNGEFLDLDLPWHSFGADSVSPANTENEVEPVDARPAADRGLTTRPGSGLTN 660
|||||
Db 601 KTTVELLPVNGEFLDLDLPWHSFGADSVSPANTENEVEPVDARPAADRGLTTRPGSGLTN 660
QY 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEVDGSKNGCAIICGLMVGGVVIATVITL 720
|||||
Db 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEVDGSKNGCAIICGLMVGGVVIATVITL 720
QY 721 VMLKKKQYTSIHGGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN 770
|||||
Db 721 VMLKKKQYTSIHGGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN 770

RESULT 11
ABB78609
ID ABB78609 standard; Protein: 770 AA.
XX
AC ABB78609;
XX
XX 16-JUL-2002 (first entry)
XX
DE Human APP770 protein sequence SEQ ID NO:55.
XX
KW Human; Asp-1; Asp-2; aspartyl protease; Alzheimer's disease;
KW proteolytic; amyloid precursor protein; APP.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX FH Misc-difference 358 /note= "encoded by GGC"
XX FT
XX
XX GB2367060-A.
XX
XX 27-MAR-2002.
XX
PD

```

XX PF 29-OCT-2001; 2001GB-0025934.
 XX PR 23-SEP-1999; 99US-155493P.
 XX PR 23-SEP-1999; 99US-0404133.
 XX PR 23-SEP-1999; 99US-0404133.
 XX PR 13-OCT-1999; 99US-0416901.
 XX PR 06-DEC-1999; 99US-169232P.
 XX PR 22-SEP-2000; 2000GB-0023315.
 XX PA (PHAA) PHARMACIA & UPJOHN CO.
 XX PI Blenkowski MJ, Gurney M;
 XX DR WPI; 2002-396337/43.
 XX DR N-PSDB; ABL52489.
 PT Human aspartyl protease 1 substrates useful in assays to detect
 PT aspartyl protease activity, e.g. for the diagnosis of Alzheimer's
 PT disease -
 XX PS Disclosure; Page 142-144; 182pp; English.
 XX CC The present invention describes a human aspartyl protease 1 (hu-Asp1)
 CC substrate (I) which comprises a peptide of no more than 50 amino acids,
 CC and which comprises the 8 amino acid sequence Gly-Leu-Ala-Leu-Ala-Leu-
 CC Glu-Pro. Also described are: (1) a method (II) for assaying hu-Asp1
 CC proteolytic activity, comprising: (a) contacting a hu-Asp1 protein with
 CC (1) under acidic conditions; and (b) determining the level of hu-Asp1
 CC proteolytic activity; (2) a purified polynucleotide (III) comprising a
 CC nucleotide sequence that hybridises under stringent conditions to the
 CC non-coding strand complementary to a defined 1804 nucleotide sequence
 CC (see ABL52456) where the nucleotide sequence encodes a polypeptide having
 CC Asp1 proteolytic activity and lacks nucleotides encoding a transmembrane
 CC domain; (3) a purified polynucleotide (III') comprising a sequence that
 CC hybridises under stringent conditions to (III) (the nucleotide sequence
 CC encodes a polypeptide further lacking a pro-peptide domain corresponding
 CC to amino acids 23-62 of hu-Asp1 (see ABB78589)); (4) a vector (IV)
 CC comprising (III) or (III'); and (5) a host cell (V) transformed or
 CC transfected with (III), (III') and/or (IV). The hu-Asp1 protease
 CC substrate (I) may be used as an enzyme substrate in assays to detect
 CC aspartyl protease activity, (II) and therefore diagnose diseases
 CC associated with aberrant hu-Asp1 expression and activity such as
 CC Alzheimer's disease. Hu-Asp1 has been localised to chromosome 21, while
 CC hu-Asp2 has been localised to chromosome 11q23.3-24.1. The present
 CC sequence represents human amyloid precursor protein APP770, which is used
 CC in the exemplification of the present invention.
 XX SQ Sequence 770 AA;
 Query Match 100.0%; Score 4058; DB 23; Length 770;
 Best Local Similarity 100.0%; Pred. No. 5.3e-289;
 Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MLPGLALLLAATARALETVDGNDAGLLAPQIATMFCGRNLNMMHNVQNGKWDSDPSGK 60
 DB 1 MLPGLALLLAATARALETVDGNDAGLLAPQIATMFCGRNLNMMHNVQNGKWDSDPSGK 60
 QY 61 TCIDFKGILQYCORVYPELOITNVVEANQPTVQNCCKRGKCKOCTHPHFVYRCLVG 120
 DB 61 TCIDFKGILQYCORVYPELOITNVVEANQPTVQNCCKRGKCKOCTHPHFVYRCLVG 120
 QY 121 EFPVSALLVPCKFLHQRMDVCEETHLHHTVAKETCSKSTNLHDYGMLLPCGIDKFR 180
 DB 121 EFPVSALLVPCKFLHQRMDVCEETHLHHTVAKETCSKSTNLHDYGMLLPCGIDKFR 180
 QY 181 GVEFVCCPLAESDNVDSADAEDSDVWVGADTDYADGSEDKVVEVAEEVEEVEE 240
 DB 181 GVEFVCCPLAESDNVDSADAEDSDVWVGADTDYADGSEDKVVEVAEEVEEVEE 240
 QY 241 EADDDEDDGDEVEEAEPEEATERTTATTTTTTTSVEEVREVSCSEQAETGPC 300
 DB 241 EADDDEDDGDEVEEAEPEEATERTTATTTTTTTSVEEVREVSCSEQAETGPC 300

QY 301 RAMISRWYFDVTEGKCAPFFYGGCGNRRNFTDEYCMVCGSAMSQSLLKTTOEPLARD 360
 DB 301 RAMISRWYFDVTEGKCAPFFYGGCGNRRNFTDEYCMVCGSAMSQSLLKTTOEPLARD 360
 QY 361 PVKLPPTAASDPADVVKYLETPGDENEHAHFOKAKERLEAKHREMSQVMREWEAEQA 420
 DB 361 PVKLPPTAASDPADVVKYLETPGDENEHAHFOKAKERLEAKHREMSQVMREWEAEQA 420
 QY 421 KNLPRADKAVIQHFQEKVESLEQEAANERQOLVETHMARVEAMLNDRRRLALENYITAL 480
 DB 421 KNLPRADKAVIQHFQEKVESLEQEAANERQOLVETHMARVEAMLNDRRRLALENYITAL 480
 QY 481 QAVPPRPFRHVFNNLKKYVRAEQKDRQHTLKHFEHVRMVDPPKAAQIRSOVMTHLRVIYER 540
 DB 481 QAVPPRPFRHVFNNLKKYVRAEQKDRQHTLKHFEHVRMVDPPKAAQIRSOVMTHLRVIYER 540
 QY 541 MNQSLLLYNPAVAEETQDEVDLLOKQNYSDOVLANMISEPRISYGNDAIMPSTLET 600
 DB 541 MNQSLLLYNPAVAEETQDEVDLLOKQNYSDOVLANMISEPRISYGNDAIMPSTLET 600
 QY 601 KTTVELLPVNGEFLDLOPHWSFGADSVPAANTENEVEPVDARPAADRGLTTRPGSLTN 660
 DB 601 KTTVELLPVNGEFLDLOPHWSFGADSVPAANTENEVEPVDARPAADRGLTTRPGSLTN 660
 QY 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVGWVIATVIVITL 720
 DB 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVGWVIATVIVITL 720
 QY 721 VMLKKQYTSIHGGVVEVDAANTPEERHLSKMQQNGYENPTYKFFEQMON 770
 DB 721 VMLKKQYTSIHGGVVEVDAANTPEERHLSKMQQNGYENPTYKFFEQMON 770
 RESULT 12
 AAG68317
 ID AAG68317 standard; Protein; 770 AA.
 XX AAG68317;
 AC AAG68317;
 XX 21-FEB-2002 (first entry)
 DT Human amyloid precursor protein (APP770) SEQ ID NO:5.
 DE Human; beta-amyloid; cyclin-dependent kinase inhibitor; nerve cell;
 XX Human; beta-amyloid; cyclin-dependent kinase inhibitor; nerve cell;
 KW amyloid precursor protein; APP.
 XX Homo sapiens.
 OS Homo sapiens.
 XX WO200182967-A1.
 PN 08-NOV-2001.
 PD 25-APR-2001; 2001WO-JP03555.
 PF 28-APR-2000; 2000JP-0131037.
 XX (YAMA) YAMANOUCHI PHARM CO LTD.
 PA (SUZU/) SUZUKI T.
 XX Suzuki T, Watanabe T, Kawabata S, Hachiya S;
 PI WPI; 2002-026209/03.
 DR Medicinal compositions for the treatment of dementia and Alzheimer's
 XX disease, comprise compounds that suppress beta amyloid production
 PT Disclosure; Page 39-42; 62pp; Japanese.
 PS The present invention describes medicinal compositions (I) inhibiting
 CC beta-amyloid production comprising an active component a substance that
 CC inhibits the activity of cyclin-dependent kinase (CDK). Also described
 CC are: (1) a method for screening compounds for their ability to inhibit

CC the production of beta-amyloid by contacting with beta-amyloid producing
 CC cells; and (2) screening kits. (1) have neurotropic and neuroprotective
 CC activities. (1) suppress the phosphorylation of amyloid precursor protein
 CC (APP) which is an essential step in the production of beta-amyloid. (1)
 CC can be used in the treatment and prevention of neurodegenerative diseases
 CC such as dementia and Alzheimer's disease. The present sequence represents
 CC the human APP770 protein which is used in the exemplification of the
 CC present invention.
 XX
 SQ Sequence 770 AA;

Query Match 100.0%; Score 4058; DB 23; Length 770;
 Best Local Similarity 100.0%; Pred. No. 5.3e-289;
 Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPOIAMFCGRNLNHNHNQVQNGKWDSPSGTK 60
 DB 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPOIAMFCGRNLNHNHNQVQNGKWDSPSGTK 60
 QY 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTIONWCKRGKQCKTHPHFVPIYRCLVG 120
 DB 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTIONWCKRGKQCKTHPHFVPIYRCLVG 120
 QY 121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCEKSTNLHDYGMLLPCGIDKFR 180
 DB 121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCEKSTNLHDYGMLLPCGIDKFR 180
 QY 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEEVEE 240
 DB 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEEVEE 240
 QY 241 EADDDDDGDEVEAEPEEATERTTSTATTTTTTTESEVEEVREVCSEQETGPC 300
 DB 241 EADDDDDGDEVEAEPEEATERTTSTATTTTTTTESEVEEVREVCSEQETGPC 300
 QY 301 RAMISRWYFDVTGECAPFFYGGCGGNRNFTTEYCMVCGSAMSQSLKTTQEPPLARD 360
 DB 301 RAMISRWYFDVTGECAPFFYGGCGGNRNFTTEYCMVCGSAMSQSLKTTQEPPLARD 360
 QY 361 PVKPLPTAASTPDVADVKYLETGPDNEHAHFOKAKERLEAKHRMSQVMREWEAEARQA 420
 DB 361 PVKPLPTAASTPDVADVKYLETGPDNEHAHFOKAKERLEAKHRMSQVMREWEAEARQA 420
 QY 421 KNLPRADKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
 DB 421 KNLPRADKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
 QY 481 QAVPRPRHVFNMLKKYVRAEQKDRQHTLKHFHVRVMDPKAAQIRSQVMTLRIYIER 540
 DB 481 QAVPRPRHVFNMLKKYVRAEQKDRQHTLKHFHVRVMDPKAAQIRSQVMTLRIYIER 540
 QY 541 MNQSLSLNVPVAAEEIQDEVELLQKQNYSDOVLANMISEPRISYGNALMPSLTET 600
 DB 541 MNQSLSLNVPVAAEEIQDEVELLQKQNYSDOVLANMISEPRISYGNALMPSLTET 600
 QY 601 KTTVELLPVNGEFLDLDLPWHSFGADSVPAANTENEVEFPVDARPAADRGLTTRPGSLTN 660
 DB 601 KTTVELLPVNGEFLDLDLPWHSFGADSVPAANTENEVEFPVDARPAADRGLTTRPGSLTN 660
 QY 661 IKTEEISEVKMDAEPHRDSEGEVHHQKLVFFAEADVGSNKGAIIGLMGVGVVIVITL 720
 DB 661 IKTEEISEVKMDAEPHRDSEGEVHHQKLVFFAEADVGSNKGAIIGLMGVGVVIVITL 720
 QY 721 VMLKKKQVTSIHGVEVYDAVTPPEERHLSKMQONGYENPTYKFEQMQN 770
 DB 721 VMLKKKQVTSIHGVEVYDAVTPPEERHLSKMQONGYENPTYKFEQMQN 770

RESULT 13
 AAEL10650
 ID AAEL10650 standard; Protein; 772 AA.
 XX
 AC AAEL10650;

XX 10-DEC-2001 (first entry)
 DE Human amyloid protein precursor 770-KK (APP770-KK) isoform.
 DE Human; aspartyl protease 1; Aspl1; amyloid precursor protein; APP770-KK;
 DE Alzheimer's disease; Ab; dementia; neurofibrillary tangle; gliosis;
 DE amyloid plaque; neuronal loss; proteolytic; neurotropic; neuroprotective.
 OS Homo sapiens.
 OS Synthetic.
 PN GB2357767-A.
 XX 04-JUL-2001.
 XX 22-SEP-2000; 2000GB-0023315.
 PR 23-SEP-1999; 99US-0155493.
 PR 23-SEP-1999; 99US-0404133.
 PR 23-SEP-1999; 99WO-US20881.
 PR 13-OCT-1999; 99US-0416901.
 PR 06-DEC-1999; 99US-0169232.
 XX (PHAA) PHARMACIA & UPJOHN CO.
 XX Blenkowski MJ, Gurney M;
 PI WPI; 2001-444208/48.
 DR N-PSDB; AAD17899.
 XX Polypeptide comprising fragments of human aspartyl protease with
 PT amyloid precursor protein processing activity and alpha-secretase
 PT activity, for identifying modulators useful in treating Alzheimer's
 PT disease -
 XX Disclosure; Page 148-151; 187pp; English.
 XX The patent discloses human aspartyl protease 1 (hu-Aspl) or modified
 CC Aspl proteins which lack transmembrane domain or amino terminal
 CC domain or cytoplasmic domain and retains alpha-secretase activity
 CC and amyloid protein precursor (APP) processing activity. The proteins
 CC of the invention are useful for assaying hu-Aspl alpha-secretase
 CC activity, which in turn is useful for identifying modulators of
 CC hu-Aspl alpha-secretase activity, where modulators that increase
 CC hu-Aspl alpha-secretase activity are useful for treating Alzheimer's
 CC disease (AD) which causes progressive dementia with consequent
 CC formation of amyloid plaques, neurofibrillary tangles, gliosis and
 CC neuronal loss. Hu-Aspl protease substrate is useful for assaying
 CC hu-Aspl proteolytic activity, by contacting hu-Aspl protein with
 CC the substrate under acidic conditions and determining the level of
 CC hu-Aspl proteolytic activity. The present sequence is human amyloid
 CC protein precursor 770-KK (APP770-KK) isoform. APP770-KK isoform is
 CC obtained by the addition of two Lys residues (KK motif) at the
 CC C-terminal of APP770 protein.
 XX Sequence 772 AA;

Query Match 100.0%; Score 4058; DB 22; Length 772;
 Best Local Similarity 100.0%; Pred. No. 5.3e-289;
 Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPOIAMFCGRNLNHNHNQVQNGKWDSPSGTK 60
 DB 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPOIAMFCGRNLNHNHNQVQNGKWDSPSGTK 60
 QY 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTIONWCKRGKQCKTHPHFVPIYRCLVG 120
 DB 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTIONWCKRGKQCKTHPHFVPIYRCLVG 120
 QY 121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCEKSTNLHDYGMLLPCGIDKFR 180
 DB 121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCEKSTNLHDYGMLLPCGIDKFR 180

QY 181 GVEFVCCPLAEEEDNDVSDADREDDSDVWVGADTDYADGSEDKVVEVAEEVEE 240
Db 181 GVEFVCCPLAEEEDNDVSDADREDDSDVWVGADTDYADGSEDKVVEVAEEVEE 240
QY 241 EADDDDEDDGDEVEEAEPEEATERTSTATTTTTTSTESVEEVVRCSEQAETGPC 300
Db 241 EADDDDEDDGDEVEEAEPEEATERTSTATTTTTTSTESVEEVVRCSEQAETGPC 300
QY 301 RAMISRWYFDVTGKCAPEFYGGCGGNRNFDTEECYMAVCGSAMQSLLKTTQEPPLARD 360
Db 301 RAMISRWYFDVTGKCAPEFYGGCGGNRNFDTEECYMAVCGSAMQSLLKTTQEPPLARD 360
QY 361 PVKLPPTAASTPDVADVKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAEEROA 420
Db 361 PVKLPPTAASTPDVADVKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAEEROA 420
QY 421 KNLPKADKKAVIQHFOEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
Db 421 KNLPKADKKAVIQHFOEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
QY 481 QAVPPRPRHVNMLKYYVRAEQDRQHTLKHFEHVRMVDPKKAAQIRSOVMTHLRVIYER 540
Db 481 QAVPPRPRHVNMLKYYVRAEQDRQHTLKHFEHVRMVDPKKAAQIRSOVMTHLRVIYER 540
QY 541 MNQSLSLYNVPAVEEIQDEVELLQKEQNYSDVLANNISPRISYGNDAIMPSTET 600
Db 541 MNQSLSLYNVPAVEEIQDEVELLQKEQNYSDVLANNISPRISYGNDAIMPSTET 600
QY 601 KTTVELLPVNGEFLSDLPQWHSFGADSVDPANTENEVEPVDARPAADRLTTRPGSGLTN 660
Db 601 KTTVELLPVNGEFLSDLPQWHSFGADSVDPANTENEVEPVDARPAADRLTTRPGSGLTN 660
QY 661 IKTEEISEVKMDAEFRHDSGYEVHQQKLVFAEDVGSNKGAIILGLMVGVIATVITL 720
Db 661 IKTEEISEVKMDAEFRHDSGYEVHQQKLVFAEDVGSNKGAIILGLMVGVIATVITL 720
QY 721 VMLKKQYTSIHGCVVEVDAVTPPEERHLSKMQONGYENPTYKFFEQMQN 770
Db 721 VMLKKQYTSIHGCVVEVDAVTPPEERHLSKMQONGYENPTYKFFEQMQN 770
RESULT 14
AAE06895
ID .AAE06895 standard; Protein; 772 AA.
AC AAE06895;
XX 23-OCT-2001 (first entry)
XX Human amyloid precursor protein 770-KK (App770-KK) isoform.
XX
XX
XX Human; aspartyl protease; Asp; beta-amyloid precursor protein 770-KK;
KW beta-secretase; Alzheimer's disease; dementia; amyloid plaque; gliosis;
KW neurofibrillary tangle; neuronal loss; amyloid-beta peptide; nontropic;
KW neuroprotective; antisense therapy; gene therapy; App770-KK; mutant;
KW mutin.
XX
XX Homo sapiens.
OS Synthetic.
OS
XX WO200150829-A2.
XX
XX 19-JUL-2001.
XX
XX 09-MAY-2001; 2001WO-IB00799.
XX
XX 09-MAY-2001; 2001WO-IB00799.
XX
XX (BIEN/) BIENKOWSKI M J.
PA (GURN/) GURNEY M E.
PA (HEIN/) HEINRIKSON R L.
PA (PARO/) PARODI L A.

(YANR/) YAN R.
XX Bienkowski MJ, Gurney ME, Heinrikson RL, Parodi LA, Yan R;
XX WPI; 2001-483072/52.
DR N-PSDB; AAD13280.
XX
XX Novel purified polypeptide comprising fragment of mammalian aspartyl
PT protease 2, lacking Asp2 transmembrane domain and retaining beta
PT secretase activity of Asp2 useful for identifying inhibitors of Asp2
PT activity
XX
XX Disclosure; Page 177-180; 185pp; English.
XX
XX The invention relates to human aspartyl proteases (Hu-Asp), beta-amyloid
CC precursor protein (APP) isoforms and their corresponding DNA molecules.
CC Human aspartyl proteases can act as beta-secretase proteases useful for
CC treating Alzheimer's disease. APP isoforms are useful for identifying
CC modulators of amyloid-beta peptide production, for use in designing
CC therapeutics for the treatment and prevention of Alzheimer's disease,
CC dementia, formation of amyloid plaques, neurofibrillary tangles, gliosis
CC and neuronal loss. APP isoforms are also used in methods for identifying
CC inhibitors and modulators of human Asp2 activity. The invention relates
CC to a method for identifying agents that modulate the activity of human
CC aspartyl protease Asp2. Amyloid-beta peptides obtained from APP are used
CC as a means to screen in cellular assays for the inhibitors of beta- and
CC gamma- secretase. Hu-Asp DNA fragments are useful as probes or primers in
CC polymerase chain reactions (PCR). The probes are useful for detecting
CC Hu-Asp nucleic acids in in vitro assays and in Northern and Southern
CC blots. The present sequence is modified human amyloid precursor protein
CC 770-KK (APP770-KK) isoform. APP770-KK isoform is obtained by addition of
CC two Lys residues (KK motif) at the C-terminal end of APP770 isoform.
XX
SQ Sequence 772 AA:
Query Match 100.0%; Score 4058; DB 22; Length 772;
Best Local Similarity 100.0%; Pred. No. 5.3e-289;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MLPGLALLLLAATAWATALEVPDGNAGLLAEPIAMFCGRLNMHNVQNGKWDSPSGTK 60
Db 1 MLPGLALLLLAATAWATALEVPDGNAGLLAEPIAMFCGRLNMHNVQNGKWDSPSGTK 60
QY 61 TCIDTKEGILQYCEQYVPELQITNVNEANQPVITONWCKGRKCKQKTHPHVIPRCILVG 120
Db 61 TCIDTKEGILQYCEQYVPELQITNVNEANQPVITONWCKGRKCKQKTHPHVIPRCILVG 120
QY 121 EFVSDALLVPDKCKFLHQRMDVCEVTHLHWHVTAKECTSEKSTNLHDYGMLLPCCIDKFR 180
Db 121 EFVSDALLVPDKCKFLHQRMDVCEVTHLHWHVTAKECTSEKSTNLHDYGMLLPCCIDKFR 180
QY 181 GVEFVCCPLAEEEDNDVSDADREDDSDVWVGADTDYADGSEDKVVEVAEEVEE 240
Db 181 GVEFVCCPLAEEEDNDVSDADREDDSDVWVGADTDYADGSEDKVVEVAEEVEE 240
QY 241 EADDDDEDDGDEVEEAEPEEATERTSTATTTTTTSTESVEEVVRCSEQAETGPC 300
Db 241 EADDDDEDDGDEVEEAEPEEATERTSTATTTTTTSTESVEEVVRCSEQAETGPC 300
QY 301 RAMISRWYFDVTGKCAPEFYGGCGGNRNFDTEECYMAVCGSAMQSLLKTTQEPPLARD 360
Db 301 RAMISRWYFDVTGKCAPEFYGGCGGNRNFDTEECYMAVCGSAMQSLLKTTQEPPLARD 360
QY 361 PVKLPPTAASTPDVADVKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAEEROA 420
Db 361 PVKLPPTAASTPDVADVKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAEEROA 420
QY 421 KNLPKADKKAVIQHFOEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
Db 421 KNLPKADKKAVIQHFOEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
QY 481 QAVPPRPRHVNMLKYYVRAEQDRQHTLKHFEHVRMVDPKKAAQIRSOVMTHLRVIYER 540
Db 481 QAVPPRPRHVNMLKYYVRAEQDRQHTLKHFEHVRMVDPKKAAQIRSOVMTHLRVIYER 540

Db 481 QAVPPRRHVFNMMLKYYVRAEQDKROHTLKHFEHVRMVDPKAAQIRSQVWTHLRIYER 540
 QY 541 MNQSLSLLYNPVAAEIQDEVDLLOKEQNYSDVLANMISEPRISYGNALMPSLTET 600
 Db 541 MNQSLSLLYNPVAAEIQDEVDLLOKEQNYSDVLANMISEPRISYGNALMPSLTET 600
 QY 601 KTTVELLPVNGEFLDLPQWHSFGADSVPAANTENEVEPVDARPAADRGLTTRPGSGLTN 660
 Db 601 KTTVELLPVNGEFLDLPQWHSFGADSVPAANTENEVEPVDARPAADRGLTTRPGSGLTN 660
 QY 661 IKTEEISEVMKDAEFRHDSYEVHHQKLVFAEDVGSNGKAIIGLMVGGVVIATVITL 720
 Db 661 IKTEEISEVMKDAEFRHDSYEVHHQKLVFAEDVGSNGKAIIGLMVGGVVIATVITL 720
 QY 721 VMLKKKYTSHHGVEVDAVTPPEERHLSKMQONGYENPTYKFFQMQN 770
 Db 721 VMLKKKYTSHHGVEVDAVTPPEERHLSKMQONGYENPTYKFFQMQN 770

RESULT 15

AAU06624

ID AAU06624 standard; Protein; 772 AA.

XX AAU06624;

XX 24-OCT-2001 (first entry)

XX Human Amyloid precursor protein mutant, APP770-KK.

XX Human: Aspartyl protease; Asp2b; beta-secretase; nootropic;
 KW neuroprotective; amyloid protein precursor; APP; Alzheimer's disease;
 KW amyloid-beta; Abeta; APP770-KK; mutant; mutain.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 358

FT /note= "Encoded by GCC"

FT Misc-difference 771..772

 FT /note= "2 Extra Lys residues added compared to
 wild-type APP770"

XX W0200149098-A2.

XX 12-JUL-2001.

XX 09-MAY-2001; 2001WO-IB00798.

XX 09-MAY-2001; 2001WO-IB00798.

XX (BIEN/) BIENKOWSKI M J.

XX (GURN/) GURNEY M E.

XX (HEIN/) HEINRIKSON R L.

XX (PARO/) PARODI L A.

XX (YANK/) YAN R.

XX Blenkowski MJ, Gurney ME, Heinrichson RL, Parodi LA, Yan R;

XX WPI; 2001-502549/55.

XX Novel purified polypeptide comprising fragment of mammalian aspartyl
 PT protease 2, lacking Asp2 transmembrane domain and retaining beta
 PT secretase activity of Asp2 useful for identifying inhibitors of Asp2
 PT activity

XX Disclosure; Page 177-180; 185pp; English.

XX The invention relates to a purified polypeptide comprising a fragment of
 CC mammalian aspartyl protease (Asp2) protein which lacks the Asp2
 CC transmembrane domain and the Asp2 protein, and where the polypeptide and
 CC the fragment retain the beta-secretase activity of the mammalian Asp2
 CC protein. The invention also details polynucleotides for the Asp
 CC proteins and vectors expressing them, and a polypeptide (isoform of

CC amyloid protein precursor (APP) comprising the amino acid sequence of an
 CC APP or its fragment containing an APP cleavage site recognizable by a
 CC mammalian beta-secretase, and further comprising two lysine residues at
 CC the carboxyl terminus of the amino acid sequence of the mammalian APP or
 CC APP fragment. Also included in the invention are methods of identifying
 CC modulators or inhibitors of Asp2. Modulators and inhibitors of Asp2 are
 CC useful for treating Alzheimer's disease. APP is useful in methods for
 CC identifying inhibitors or modulators of human Asp2 activity and
 CC amyloid-beta (Abeta) peptide production. APP is also useful in designing
 CC therapeutics for the treatment or prevention of Alzheimer's disease,
 CC APP comprising the APP-sw-beta-secretase peptide sequence (NLDA), which
 CC is associated with increased levels of Abeta processing is useful in
 CC assays relating the Alzheimer's research. The expression vector is useful
 CC for recombinantly expressing APP. Nucleic acids that hybridize to
 CC Asp oligonucleotides are useful as probes or primers. The probes are
 CC useful for detecting Hu-Asp nucleic acids in in vitro assays and in
 CC Northern and Southern blots. The present sequence is the human
 CC APP770 mutant, APP770-KK which has 2 extra Lys residues added at
 CC the C-terminus compared to the wild-type APP770. The mutation alters the
 CC specificity of the APP gamma-secretase activity and increases the rate
 CC of processing of the amyloid Abeta peptide.

XX Sequence 772 AA;

Query Match 100.0%; Score 4058; DB 22; Length 772;
 Best Local Similarity 100.0%; Pred. No. 5.3e-289;
 Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPOIAMFCGRLNMHNVONGKWDSPSGTK 60
 Db 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPOIAMFCGRLNMHNVONGKWDSPSGTK 60

QY 61 TCIDTKEGILQYCEVYPPELQITNVVEANQPVPTIQNWKCKRQCKTTHPHFVYRCLVG 120
 Db 61 TCIDTKEGILQYCEVYPPELQITNVVEANQPVPTIQNWKCKRQCKTTHPHFVYRCLVG 120

QY 121 EFVSDALLVPDKCKFLHOERMDVCETHLHWHVTAKETSEKSTNLHDYGMLLPCGIDKFR 180
 Db 121 EFVSDALLVPDKCKFLHOERMDVCETHLHWHVTAKETSEKSTNLHDYGMLLPCGIDKFR 180

QY 181 GVEFYCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEEVEEVEE 240
 Db 181 GVEFYCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEEVEEVEE 240

QY 241 EADDEDEDDGEVEEAEPEYEATERTTSIATTTTTTTSVEEVREVCSQAETGPC 300
 Db 241 EADDEDEDDGEVEEAEPEYEATERTTSIATTTTTTTSVEEVREVCSQAETGPC 300

QY 301 RAMISRWFYDVTGKCAPFFYGGCGGNRNNEDTEEYCHAVCGSAMQSLLKTTQEPPLARD 360
 Db 301 RAMISRWFYDVTGKCAPFFYGGCGGNRNNEDTEEYCHAVCGSAMQSLLKTTQEPPLARD 360

QY 361 PVKLPTTAASTPDVADKYLETTPGDENEHAHFQKAKERLEAKHRRMSQVMREWEAEERQA 420
 Db 361 PVKLPTTAASTPDVADKYLETTPGDENEHAHFQKAKERLEAKHRRMSQVMREWEAEERQA 420

QY 421 KNLPRADKAVIQHFOEKVESLEQEAANERQOLVETHMARVEAMLNDRRRRLALENYITAL 480
 Db 421 KNLPRADKAVIQHFOEKVESLEQEAANERQOLVETHMARVEAMLNDRRRRLALENYITAL 480

QY 481 QAVPPRRHVFNMMLKYYVRAEQDKROHTLKHFEHVRMVDPKAAQIRSQVWTHLRIYER 540
 Db 481 QAVPPRRHVFNMMLKYYVRAEQDKROHTLKHFEHVRMVDPKAAQIRSQVWTHLRIYER 540

QY 541 MNQSLSLLYNPVAAEIQDEVDLLOKEQNYSDVLANMISEPRISYGNALMPSLTET 600
 Db 541 MNQSLSLLYNPVAAEIQDEVDLLOKEQNYSDVLANMISEPRISYGNALMPSLTET 600

QY 601 KTTVELLPVNGEFLDLPQWHSFGADSVPAANTENEVEPVDARPAADRGLTTRPGSGLTN 660
 Db 601 KTTVELLPVNGEFLDLPQWHSFGADSVPAANTENEVEPVDARPAADRGLTTRPGSGLTN 660

QY 661 IKTEEISEVMKDAEFRHDSYEVHHQKLVFAEDVGSNGKAIIGLMVGGVVIATVITL 720

|||||
Db 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIAVITL 720
QY 721 VMLKKKOYTSIHGCVVEVDAAVTPEERHLSKMOONGYENPTYKFFEQMN 770
Db 721 VMLKKKOYTSIHGCVVEVDAAVTPEERHLSKMOONGYENPTYKFFEQMN 770

Search completed: January 28, 2003, 14:13:55
Job time : 45.9063 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:19:14 ; Search time 10.0956 Seconds
(without alignments)
1539.035 Million cell updates/sec

Title: US-09-904-987-2
Perfect score: 4058
Sequence: 1 MLPGLALLLAANTARALEV.....KMQQGYENPTYKFEQMON 770

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 122226 seqs, 20178551 residues
Total number of hits satisfying chosen parameters: 122226

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4058	100.0	770	9	US-09-785-215-2
2	4058	100.0	770	10	US-09-794-927-55
3	4058	100.0	770	10	US-09-795-847-55
4	4058	100.0	770	10	US-09-794-743-55
5	4058	100.0	770	10	US-09-794-748-55
6	4058	100.0	770	10	US-09-904-987-2
7	4058	100.0	770	10	US-09-794-925-55
8	4058	100.0	770	10	US-09-681-442-55
9	4058	100.0	770	10	US-09-149-718-6
10	4058	100.0	772	10	US-09-794-927-59
11	4058	100.0	772	10	US-09-795-847-59
12	4058	100.0	772	10	US-09-794-743-59
13	4058	100.0	772	10	US-09-794-748-59
14	4058	100.0	772	10	US-09-794-925-59
15	4058	100.0	772	10	US-09-681-442-59
16	3943.5	97.2	751	10	US-09-794-927-57
17	3943.5	97.2	751	10	US-09-795-847-57
18	3943.5	97.2	751	10	US-09-794-743-57
19	3943.5	97.2	751	10	US-09-794-748-57

20	3943.5	97.2	751	10	US-09-794-925-57	Sequence 57, Appl
21	3943.5	97.2	751	10	US-09-681-442-57	Sequence 57, Appl
22	3943.5	97.2	751	10	US-09-149-718-4	Sequence 4, Appl
23	3943.5	97.2	753	10	US-09-794-927-61	Sequence 61, Appl
24	3943.5	97.2	753	10	US-09-795-847-61	Sequence 61, Appl
25	3943.5	97.2	753	10	US-09-794-743-61	Sequence 61, Appl
26	3943.5	97.2	753	10	US-09-794-748-61	Sequence 61, Appl
27	3943.5	97.2	753	10	US-09-794-925-61	Sequence 61, Appl
28	3943.5	97.2	753	10	US-09-681-442-61	Sequence 61, Appl
29	3590.5	88.5	695	10	US-09-794-927-10	Sequence 10, Appl
30	3590.5	88.5	695	10	US-09-795-847-10	Sequence 10, Appl
31	3590.5	88.5	695	10	US-09-794-743-10	Sequence 10, Appl
32	3590.5	88.5	695	10	US-09-794-748-10	Sequence 10, Appl
33	3590.5	88.5	695	10	US-09-794-925-10	Sequence 10, Appl
34	3590.5	88.5	695	10	US-09-681-442-10	Sequence 10, Appl
35	3590.5	88.5	695	10	US-09-149-718-2	Sequence 2, Appl
36	3590.5	88.5	697	10	US-09-794-927-16	Sequence 16, Appl
37	3590.5	88.5	697	10	US-09-795-847-16	Sequence 16, Appl
38	3590.5	88.5	697	10	US-09-794-743-16	Sequence 16, Appl
39	3590.5	88.5	697	10	US-09-794-748-16	Sequence 16, Appl
40	3590.5	88.5	697	10	US-09-794-925-16	Sequence 16, Appl
41	3590.5	88.5	697	10	US-09-681-442-16	Sequence 16, Appl
42	3585.5	88.4	695	10	US-09-794-927-14	Sequence 14, Appl
43	3585.5	88.4	695	10	US-09-795-847-14	Sequence 14, Appl
44	3585.5	88.4	695	10	US-09-794-743-14	Sequence 14, Appl
45	3585.5	88.4	695	10	US-09-794-748-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1
US-09-785-215-2
; Sequence 2, Application US/09785215
; Publication No. US20020187157A1
; GENERAL INFORMATION:
; APPLICANT: JENSEN, Martin Roland et al.
; TITLE OF INVENTION: NOVEL METHOD FOR DOWN-REGULATION OF AMYLOID
; FILE REFERENCE: 3631-0107P
; CURRENT APPLICATION NUMBER: US/09/785,215
; CURRENT FILING DATE: 2001-02-20
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 770
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-785-215-2

Query Match 100.0%; Score 4058; DB 9; Length 770;
Best Local Similarity 100.0%; Pred. No. 3.9e-267;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MLPGLALLLAANTARALEVPTDGNAGLLAEPQIATMFCGRNLHMNVQNGKWDSPSGTK	60
Db	1	MLPGLALLLAANTARALEVPTDGNAGLLAEPQIATMFCGRNLHMNVQNGKWDSPSGTK	60
QY	61	TCIDTKEGILQYCOEYVPELQITNVVEANQPVTIQNCKRGRKQCKTHPHFVYPRCLVG	120
Db	61	TCIDTKEGILQYCOEYVPELQITNVVEANQPVTIQNCKRGRKQCKTHPHFVYPRCLVG	120
QY	121	EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSKSTNLHDYGMLLPCGIDKFR	180
Db	121	EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSKSTNLHDYGMLLPCGIDKFR	180
QY	181	GVEFVCCPLAESDNVDSADAEEDSDVMWGGADTDYADGSEDKVVEAEVEEVEE	240
Db	181	GVEFVCCPLAESDNVDSADAEEDSDVMWGGADTDYADGSEDKVVEAEVEEVEE	240
QY	241	EADDEDEDEGDEVEEAEPYEATERTTTSIATTTTTTTSVEEVVREVCSEQAETGPC	300
Db	241	EADDEDEDEGDEVEEAEPYEATERTTTSIATTTTTTTSVEEVVREVCSEQAETGPC	300

;; PRIOR APPLICATION NUMBER: PCT/US99/20881
;; PRIOR FILING DATE: 1999-09-23
;; PRIOR APPLICATION NUMBER: 60/101,594
;; PRIOR FILING DATE: 1998-09-24
;; NUMBER OF SEQ ID NOS: 73
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 55
;; LENGTH: 770
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-795-847-55

Query Match 100.0%; Score 4058; DB 10; Length 770;
Best Local Similarity 100.0%; Pred. No. 3.9e-267;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MLPGLALLLAATAARALEVPTDGNAGLLAEPQIAMFCGRLLNMHNQVNGKWDSPSGTK 60
DB 1 MLPGLALLLAATAARALEVPTDGNAGLLAEPQIAMFCGRLLNMHNQVNGKWDSPSGTK 60
QY 61 TCIDTREGILQYCOEYVPELQITNVVEANQPVTIONCKRGRKCKTHPHFVPIYRCLVG 120
DB 61 TCIDTREGILQYCOEYVPELQITNVVEANQPVTIONCKRGRKCKTHPHFVPIYRCLVG 120
QY 121 EFVSDALLVPDKCFKFLHQMMDVCETHLHWHTVAKETCEKSTNLHDYGMLLPCGIDKFR 180
DB 121 EFVSDALLVPDKCFKFLHQMMDVCETHLHWHTVAKETCEKSTNLHDYGMLLPCGIDKFR 180
QY 181 GVEFVCCPLAESDNVDSADAEEDSDVWGGADTDYADGSEDKVVEVAEEVAEVEE 240
DB 181 GVEFVCCPLAESDNVDSADAEEDSDVWGGADTDYADGSEDKVVEVAEEVAEVEE 240
QY 241 EADDDDEDDGDEVEEAEPEEATERTTSIATTTTTTSTESVEEVVREVCSEQAETGPC 300
DB 241 EADDDDEDDGDEVEEAEPEEATERTTSIATTTTTTSTESVEEVVREVCSEQAETGPC 300
QY 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNFDTEEYCMVCGSAMSQSLLKTTQPEPLARD 360
DB 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNFDTEEYCMVCGSAMSQSLLKTTQPEPLARD 360
QY 361 PVKLPPTAASTPDVADKYLETGPDNEHAHFQKAKERLEAKHRMSQVMEWEAEARQA 420
DB 361 PVKLPPTAASTPDVADKYLETGPDNEHAHFQKAKERLEAKHRMSQVMEWEAEARQA 420
QY 421 KNLPKADKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
DB 421 KNLPKADKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
QY 481 QAVPPRPRHVFNMKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIER 540
DB 481 QAVPPRPRHVFNMKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIER 540
QY 541 MNQSLSLLYNVPAAVEIQDEVDLQKEQNYSDVLANMISEPRISYGNDAIMPSTLET 600
DB 541 MNQSLSLLYNVPAAVEIQDEVDLQKEQNYSDVLANMISEPRISYGNDAIMPSTLET 600
QY 601 KTTVELLPVNGEFLDLDLPWHISFGADSVPAANTEVEPVPDARPAADRLTTPGSGGLTN 660
DB 601 KTTVELLPVNGEFLDLDLPWHISFGADSVPAANTEVEPVPDARPAADRLTTPGSGGLTN 660
QY 661 IKTEEISEVKMDAEPHDSGYEVHQQLVFFPAEDVGSNGKGAIGLMVGGVVIATVITL 720
DB 661 IKTEEISEVKMDAEPHDSGYEVHQQLVFFPAEDVGSNGKGAIGLMVGGVVIATVITL 720
QY 721 VMLKKKQVTSIHGGVVEVDAVTPPEERHLSKMQONGYENPTYKFEQMON 770
DB 721 VMLKKKQVTSIHGGVVEVDAVTPPEERHLSKMQONGYENPTYKFEQMON 770
```

RESULT 4
US-09-794-743-55
; Sequence 55, Application US/09794743
; Patent No. US20010021391A1

;; GENERAL INFORMATION:
;; APPLICANT: Gurney, Mark E.
;; APPLICANT: Blenkowski, Michael J.
;; APPLICANT: Heinrikson, Robert L.
;; APPLICANT: Parodi, Luis A.
;; APPLICANT: Yan, Riqiang
;; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
;; TITLE OF INVENTION: USES
;; TITLE OF INVENTION: THEREFOR
;; FILE REFERENCE: 28341/6280BC
;; CURRENT APPLICATION NUMBER: US/09/794,743
;; CURRENT FILING DATE: 2001-02-27
;; PRIOR APPLICATION NUMBER: 09/416,901
;; PRIOR FILING DATE: 1999-10-13
;; PRIOR APPLICATION NUMBER: 60/155,493
;; PRIOR FILING DATE: 1999-09-23
;; PRIOR APPLICATION NUMBER: 09/404,133
;; PRIOR FILING DATE: 1999-09-23
;; PRIOR APPLICATION NUMBER: PCT/US99/20881
;; PRIOR FILING DATE: 1999-09-23
;; PRIOR APPLICATION NUMBER: 60/101,594
;; PRIOR FILING DATE: 1998-09-24
;; NUMBER OF SEQ ID NOS: 73
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 55
;; LENGTH: 770
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-794-743-55

Query Match 100.0%; Score 4058; DB 10; Length 770;
Best Local Similarity 100.0%; Pred. No. 3.9e-267;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MLPGLALLLAATAARALEVPTDGNAGLLAEPQIAMFCGRLLNMHNQVNGKWDSPSGTK 60
DB 1 MLPGLALLLAATAARALEVPTDGNAGLLAEPQIAMFCGRLLNMHNQVNGKWDSPSGTK 60
QY 61 TCIDTREGILQYCOEYVPELQITNVVEANQPVTIONCKRGRKCKTHPHFVPIYRCLVG 120
DB 61 TCIDTREGILQYCOEYVPELQITNVVEANQPVTIONCKRGRKCKTHPHFVPIYRCLVG 120
QY 121 EFVSDALLVPDKCFKFLHQMMDVCETHLHWHTVAKETCEKSTNLHDYGMLLPCGIDKFR 180
DB 121 EFVSDALLVPDKCFKFLHQMMDVCETHLHWHTVAKETCEKSTNLHDYGMLLPCGIDKFR 180
QY 181 GVEFVCCPLAESDNVDSADAEEDSDVWGGADTDYADGSEDKVVEVAEEVAEVEE 240
DB 181 GVEFVCCPLAESDNVDSADAEEDSDVWGGADTDYADGSEDKVVEVAEEVAEVEE 240
QY 241 EADDDDEDDGDEVEEAEPEEATERTTSIATTTTTTSTESVEEVVREVCSEQAETGPC 300
DB 241 EADDDDEDDGDEVEEAEPEEATERTTSIATTTTTTSTESVEEVVREVCSEQAETGPC 300
QY 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNFDTEEYCMVCGSAMSQSLLKTTQPEPLARD 360
DB 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNFDTEEYCMVCGSAMSQSLLKTTQPEPLARD 360
QY 361 PVKLPPTAASTPDVADKYLETGPDNEHAHFQKAKERLEAKHRMSQVMEWEAEARQA 420
DB 361 PVKLPPTAASTPDVADKYLETGPDNEHAHFQKAKERLEAKHRMSQVMEWEAEARQA 420
QY 421 KNLPKADKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
DB 421 KNLPKADKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
QY 481 QAVPPRPRHVFNMKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIER 540
DB 481 QAVPPRPRHVFNMKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIER 540
QY 541 MNQSLSLLYNVPAAVEIQDEVDLQKEQNYSDVLANMISEPRISYGNDAIMPSTLET 600
DB 541 MNQSLSLLYNVPAAVEIQDEVDLQKEQNYSDVLANMISEPRISYGNDAIMPSTLET 600
```

```
QY 601 KTTVELLPVNGEFLDLPQWHSFGADSVDPANTENEVEPVDARPAADRLGTLTRPGSGLTN 660
|||||
Db 601 KTTVELLPVNGEFLDLPQWHSFGADSVDPANTENEVEPVDARPAADRLGTLTRPGSGLTN 660
|||||
QY 661 IKTEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNGKAIIGLMVGGVVIATVITL 720
|||||
Db 661 IKTEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNGKAIIGLMVGGVVIATVITL 720
|||||
QY 721 VMLKKQYTSIHGVEVDAAVTPPEERHLSKMQONGYENPTYKFFEQMON 770
|||||
Db 721 VMLKKQYTSIHGVEVDAAVTPPEERHLSKMQONGYENPTYKFFEQMON 770
|||||

RESULT 5
US-09-794-748-55
; Sequence 55, Application US/09794748
; Patent No. US20020037315A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Heinrikson, Robert L.
; APPLICANT: Parodi, Luis A.
; APPLICANT: Yan, Riqiang
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
; TITLE OF INVENTION: USES
; FILE REFERENCE: 28341/5280JL
; CURRENT APPLICATION NUMBER: US/09/794,748
; PRIOR FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 09/416,901
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: 60/155,493
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 55
; LENGTH: 770
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-794-748-55

Query Match 100.0%; Score 4058; DB 10; Length 770;
Best Local Similarity 100.0%; Pred. No. 3.9e-267;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLPGLALLLAAWTARALEVPTDGNAGLLAEPOIAMFCGRLNMNMVQNGKWDSPSGTK 60
|||||
Db 1 MLPGLALLLAAWTARALEVPTDGNAGLLAEPOIAMFCGRLNMNMVQNGKWDSPSGTK 60
|||||
QY 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVPIYRCLVG 120
|||||
Db 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVPIYRCLVG 120
|||||
QY 121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
|||||
Db 121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
|||||
QY 181 GVEFVCCPLAESDNVDSADAEEDSDVWGGADTDYADGSEDKVVEAEVEEVEE 240
|||||
Db 181 GVEFVCCPLAESDNVDSADAEEDSDVWGGADTDYADGSEDKVVEAEVEEVEE 240
|||||
QY 241 EADDEDDEDGEVEEAEPEEATERTTTSIATTTTTTSEVVEVREYVCSQAETGPC 300
|||||
Db 241 EADDEDDEDGEVEEAEPEEATERTTTSIATTTTTTSEVVEVREYVCSQAETGPC 300
|||||
QY 301 RAMISRWFVDETEGKCAFFFGCGGNRNNFDTEEYCMVCGSAMSQSLKTTQEP LARD 360
|||||
```

```
Db 301 RAMISRWFVDETEGKCAFFFGCGGNRNNFDTEEYCMVCGSAMSQSLKTTQEP LARD 360
|||||
QY 361 PVKLPPTTAASPTDAVDKYLETPTGDENEHAHFQKAKERLEAKHREMSQVMREWEAEARQA 420
|||||
Db 361 PVKLPPTTAASPTDAVDKYLETPTGDENEHAHFQKAKERLEAKHREMSQVMREWEAEARQA 420
|||||
QY 421 KNLPRADKAVIQHFQEKVESLEQEAANERQOLVETHMARVEAMLNDRRRRLALENYITAL 480
|||||
Db 421 KNLPRADKAVIQHFQEKVESLEQEAANERQOLVETHMARVEAMLNDRRRRLALENYITAL 480
|||||
QY 481 QAVPPRPRHVNMLKKYVRAEQKDRQHTLKHFEHVRVMDPKKAAQIRSQVMTHLRVIER 540
|||||
Db 481 QAVPPRPRHVNMLKKYVRAEQKDRQHTLKHFEHVRVMDPKKAAQIRSQVMTHLRVIER 540
|||||
QY 541 MNQSLSLLYNPVAEEIQDEVDLLOKEQNYSDDVLANMISEPRISYGNDAIMPSTET 600
|||||
Db 541 MNQSLSLLYNPVAEEIQDEVDLLOKEQNYSDDVLANMISEPRISYGNDAIMPSTET 600
|||||
QY 601 KTTVELLPVNGEFLDLPQWHSFGADSVDPANTENEVEPVDARPAADRLGTLTRPGSGLTN 660
|||||
Db 601 KTTVELLPVNGEFLDLPQWHSFGADSVDPANTENEVEPVDARPAADRLGTLTRPGSGLTN 660
|||||
QY 661 IKTEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNGKAIIGLMVGGVVIATVITL 720
|||||
Db 661 IKTEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNGKAIIGLMVGGVVIATVITL 720
|||||
QY 721 VMLKKQYTSIHGVEVDAAVTPPEERHLSKMQONGYENPTYKFFEQMON 770
|||||
Db 721 VMLKKQYTSIHGVEVDAAVTPPEERHLSKMQONGYENPTYKFFEQMON 770
|||||

RESULT 6
US-09-904-987-2
; Sequence 2, Application US/09504987
; Patent No. US20020037908A1
; GENERAL INFORMATION:
; APPLICANT: No. US20020037908Alactyl, Inc.
; TITLE OF INVENTION: Methods and Compositions for Controlling Pathological and Prop
; TITLE OF INVENTION: Protein Assembly or Aggregation
; FILE REFERENCE: 42108/26146
; CURRENT APPLICATION NUMBER: US/09/904,987
; CURRENT FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 2
; LENGTH: 770
; TYPE: PRT
; ORGANISM: homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NCBI ENTREZ / QRHUA4
; DATABASE ENTRY DATE: 2000-09-15
; RELEVANT RESIDUES: (1)..(770)
US-09-904-987-2

Query Match 100.0%; Score 4058; DB 10; Length 770;
Best Local Similarity 100.0%; Pred. No. 3.9e-267;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLPGLALLLAAWTARALEVPTDGNAGLLAEPOIAMFCGRLNMNMVQNGKWDSPSGTK 60
|||||
Db 1 MLPGLALLLAAWTARALEVPTDGNAGLLAEPOIAMFCGRLNMNMVQNGKWDSPSGTK 60
|||||
QY 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVPIYRCLVG 120
|||||
Db 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVPIYRCLVG 120
|||||
QY 121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
|||||
Db 121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
|||||
QY 181 GVEFVCCPLAESDNVDSADAEEDSDVWGGADTDYADGSEDKVVEAEVEEVEE 240
|||||
```

Db 184 GVEFVCCPLAESDNDSDADAEEDSDVWVGADTDYADGSEDKVVEVAEEVEE 240
QY 241 EADDEDEDGEVEEAEPYEATERTSIATTTTTTSTESVEEVREVCSQAEATGPC 300
Db 241 EADDEDEDGEVEEAEPYEATERTSIATTTTTTSTESVEEVREVCSQAEATGPC 300
QY 301 RAMISRWFVTEGKCAPFFYGGCGGNRNFDEEYCMVCGSAMSQSLLKTQEP LARD 360
Db 301 RAMISRWFVTEGKCAPFFYGGCGGNRNFDEEYCMVCGSAMSQSLLKTQEP LARD 360
QY 361 PVKLPTTAASDPDAVDKYLETPGDENEHAHFOKAKERLEAKHRMSQVMREWEAERQA 420
Db 361 PVKLPTTAASDPDAVDKYLETPGDENEHAHFOKAKERLEAKHRMSQVMREWEAERQA 420
QY 421 KNLPRADKKAIVIOHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
Db 421 KNLPRADKKAIVIOHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
QY 481 QAVPPRRPRHVNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKAAQIRSQVMTHLRVIYER 540
Db 481 QAVPPRRPRHVNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKAAQIRSQVMTHLRVIYER 540
QY 541 MNQSLSLLYNVPVAEEIQDEVDLQKEQNYSDVLANMISEPRISYNDALMPSLTET 600
Db 541 MNQSLSLLYNVPVAEEIQDEVDLQKEQNYSDVLANMISEPRISYNDALMPSLTET 600
QY 601 KTTVELLPVNGEFLDLDQPHSFGADSVPAANTEVEPVDARPAADRGLTTRPGSGLTN 660
Db 601 KTTVELLPVNGEFLDLDQPHSFGADSVPAANTEVEPVDARPAADRGLTTRPGSGLTN 660
QY 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVGWVIATVITL 720
Db 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVGWVIATVITL 720
QY 721 VMLKKKQYTSIIHGVVEVDAAVTPEERHLSKMQONGYENPTYKFFEQMQN 770
Db 721 VMLKKKQYTSIIHGVVEVDAAVTPEERHLSKMQONGYENPTYKFFEQMQN 770

RESULT 7
US-09-794-925-55
; Sequence 55, Application US/09794925
; Patent No. US20020064819A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Heinrikson, Robert L.
; APPLICANT: Parodi, Luis A.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND USES THEREFOR
; FILE REFERENCE: 28341/6280H1
; CURRENT APPLICATION NUMBER: US/09/794,925
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 09/416,901
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 55
; LENGTH: 770
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-794-925-55

Query Match 100.0%; Score 4058; DB 10; Length 770;

Best Local Similarity 100.0%; Pred. No. 3.9e-267;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLPGIALLLLAAWTARALEVPTDGNAGLLAEPOIAMFCGRLNMHNVMONGKWDSPSGTK 60
Db 1 MLPGIALLLLAAWTARALEVPTDGNAGLLAEPOIAMFCGRLNMHNVMONGKWDSPSGTK 60
QY 61 TCIDTKEGILQCYQEVYPELQITNVVEANQPVTIQNMCKRGRKQCKTHPHFVPIYRCILVG 120
Db 61 TCIDTKEGILQCYQEVYPELQITNVVEANQPVTIQNMCKRGRKQCKTHPHFVPIYRCILVG 120
QY 121 EFVSDALLVPDKCKFLHQERMDVVCETHLHWHVTAKETCSEKSTNLHDYGMLLPCGIDKFR 180
Db 121 EFVSDALLVPDKCKFLHQERMDVVCETHLHWHVTAKETCSEKSTNLHDYGMLLPCGIDKFR 180
QY 181 GVEFVCCPLAESDNDSDADAEEDSDVWVGADTDYADGSEDKVVEVAEEVEE 240
Db 181 GVEFVCCPLAESDNDSDADAEEDSDVWVGADTDYADGSEDKVVEVAEEVEE 240
QY 241 EADDEDEDGEVEEAEPYEATERTTSIATTTTTTSTESVEEVREVCSQAEATGPC 300
Db 241 EADDEDEDGEVEEAEPYEATERTTSIATTTTTTSTESVEEVREVCSQAEATGPC 300
QY 301 RAMISRWFVTEGKCAPFFYGGCGGNRNFDEEYCMVCGSAMSQSLLKTQEP LARD 360
Db 301 RAMISRWFVTEGKCAPFFYGGCGGNRNFDEEYCMVCGSAMSQSLLKTQEP LARD 360
QY 361 PVKLPTTAASDPDAVDKYLETPGDENEHAHFOKAKERLEAKHRMSQVMREWEAERQA 420
Db 361 PVKLPTTAASDPDAVDKYLETPGDENEHAHFOKAKERLEAKHRMSQVMREWEAERQA 420
QY 421 KNLPRADKKAIVIOHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
Db 421 KNLPRADKKAIVIOHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
QY 481 QAVPPRRPRHVNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKAAQIRSQVMTHLRVIYER 540
Db 481 QAVPPRRPRHVNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKAAQIRSQVMTHLRVIYER 540
QY 541 MNQSLSLLYNVPVAEEIQDEVDLQKEQNYSDVLANMISEPRISYNDALMPSLTET 600
Db 541 MNQSLSLLYNVPVAEEIQDEVDLQKEQNYSDVLANMISEPRISYNDALMPSLTET 600
QY 601 KTTVELLPVNGEFLDLDQPHSFGADSVPAANTEVEPVDARPAADRGLTTRPGSGLTN 660
Db 601 KTTVELLPVNGEFLDLDQPHSFGADSVPAANTEVEPVDARPAADRGLTTRPGSGLTN 660
QY 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVGWVIATVITL 720
Db 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVGWVIATVITL 720
QY 721 VMLKKKQYTSIIHGVVEVDAAVTPEERHLSKMQONGYENPTYKFFEQMQN 770
Db 721 VMLKKKQYTSIIHGVVEVDAAVTPEERHLSKMQONGYENPTYKFFEQMQN 770

RESULT 8

US-09-681-442-55
; Sequence 55, Application US/09681442
; Patent No. US20020081634A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Heinrikson, Robert L.
; APPLICANT: Parodi, Luis A.
; APPLICANT: Yan, Riqiang
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND USES THEREFOR
; FILE REFERENCE: 28341/6280FG
; CURRENT APPLICATION NUMBER: US/09/681,442
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 09/416,901
; PRIOR FILING DATE: 1999-10-13

```
; PRIOR APPLICATION NUMBER: 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 55
; LENGTH: 770
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-681-442-55

Query Match          100.0%; Score 4058; DB 10; Length 770;
Best Local Similarity 100.0%; Pred. No. 3.9e-267;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLPLGLALLLAATAWTAARALEVPTDGNAGLLAEPQIAMFCGRLNMHNQNGKWDSDPSGK 60
Db 1 MLPLGLALLLAATAWTAARALEVPTDGNAGLLAEPQIAMFCGRLNMHNQNGKWDSDPSGK 60

QY 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTIONWCKRGRKQCKTHPHFVPIYRCLVG 120
Db 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTIONWCKRGRKQCKTHPHFVPIYRCLVG 120

QY 121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHVTAKETCSEKSTNLHDYGMLLPCGIDKFR 180
Db 121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHVTAKETCSEKSTNLHDYGMLLPCGIDKFR 180

QY 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSDKVKVEVAEEVEEVEE 240
Db 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSDKVKVEVAEEVEEVEE 240

QY 241 EADDDDEDDGDEVEEAEPEEATERTTSIATTTTTTSTESVEEVVREVCSEQAETGPC 300
Db 241 EADDDDEDDGDEVEEAEPEEATERTTSIATTTTTTSTESVEEVVREVCSEQAETGPC 300

QY 301 RAMISRNYFDVTEGKCAPFFYGGCGGNRNFDTEECVMAVCGSAMSQSLLKTTQEPPLARD 360
Db 301 RAMISRNYFDVTEGKCAPFFYGGCGGNRNFDTEECVMAVCGSAMSQSLLKTTQEPPLARD 360

QY 361 PVKLPPTAASTPDVADKYLETGPDENEHAHFOKAKERLEAKHERMSQVMREWEAEARQA 420
Db 361 PVKLPPTAASTPDVADKYLETGPDENEHAHFOKAKERLEAKHERMSQVMREWEAEARQA 420

QY 421 KNLPKADKRAVIOHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
Db 421 KNLPKADKRAVIOHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480

QY 481 QAVPPRPRHVNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSOVMTHLRVIER 540
Db 481 QAVPPRPRHVNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSOVMTHLRVIER 540

QY 541 MNQSLSLYNVPAVAEEIQDEVDLQEKQNTSDVLANNISEPRISYGNDAIMPSTLET 600
Db 541 MNQSLSLYNVPAVAEEIQDEVDLQEKQNTSDVLANNISEPRISYGNDAIMPSTLET 600

QY 601 KTTVELLPVNGEFLDQFWSHFGADSVPAANTENEVEPVDARPAADRLGTLTRPGSLTN 660
Db 601 KTTVELLPVNGEFLDQFWSHFGADSVPAANTENEVEPVDARPAADRLGTLTRPGSLTN 660

QY 661 IKTEELSEVKMDAEFRHDSGYEVHHOKLVFFAEDVGSNKGAIIGLMVGCVIATVITL 720
Db 661 IKTEELSEVKMDAEFRHDSGYEVHHOKLVFFAEDVGSNKGAIIGLMVGCVIATVITL 720

QY 721 VMLKKQYTSIHGGVVEVDAATPEERHLSKMQONGYENPTYKFFEQMQN 770
Db 721 VMLKKQYTSIHGGVVEVDAATPEERHLSKMQONGYENPTYKFFEQMQN 770
```

```
RESULT 9
US-09-149-718-6
; Sequence 6, Application US/09149718
; Patent No. US20020104104A1
; GENERAL INFORMATION:
; APPLICANT: Dora K. Games, Dale B. Schenk, Lisa C. McConlogue,
; APPLICANT: Peter A. Seubert, and Russell E. Rydel
; TITLE OF INVENTION: Method For Identifying Alzheimer's Disease
; TITLE OF INVENTION: Therapeutics Using Transgenic Animal Models
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; CITY: Atlanta
; STATE: GA
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/149,718
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/660,487
; FILING DATE:
; APPLICATION NUMBER: 08/480,653
; FILING DATE: June 7, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: ANS101CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404)-873-8794
; TELEFAX: (404)-873-8795
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 770 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-149-718-6

Query Match          100.0%; Score 4058; DB 10; Length 770;
Best Local Similarity 100.0%; Pred. No. 3.9e-267;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLPLGLALLLAATAWTAARALEVPTDGNAGLLAEPQIAMFCGRLNMHNQNGKWDSDPSGK 60
Db 1 MLPLGLALLLAATAWTAARALEVPTDGNAGLLAEPQIAMFCGRLNMHNQNGKWDSDPSGK 60

QY 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTIONWCKRGRKQCKTHPHFVPIYRCLVG 120
Db 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTIONWCKRGRKQCKTHPHFVPIYRCLVG 120

QY 121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHVTAKETCSEKSTNLHDYGMLLPCGIDKFR 180
Db 121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHVTAKETCSEKSTNLHDYGMLLPCGIDKFR 180

QY 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSDKVKVEVAEEVEEVEE 240
Db 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSDKVKVEVAEEVEEVEE 240

QY 241 EADDDDEDDGDEVEEAEPEEATERTTSIATTTTTTSTESVEEVVREVCSEQAETGPC 300
Db 241 EADDDDEDDGDEVEEAEPEEATERTTSIATTTTTTSTESVEEVVREVCSEQAETGPC 300

QY 301 RAMISRNYFDVTEGKCAPFFYGGCGGNRNFDTEECVMAVCGSAMSQSLLKTTQEPPLARD 360
Db 301 RAMISRNYFDVTEGKCAPFFYGGCGGNRNFDTEECVMAVCGSAMSQSLLKTTQEPPLARD 360
```

```
Db 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNFDTEYCMVCGSAMQSLLKTTQEPPLARD 360
QY 361 PVKLPPTAASTPDVADKYLETGPDENEHAHFQKAKERLEAKHRMSQVMREWEAEARQA 420
Db 361 PVKLPPTAASTPDVADKYLETGPDENEHAHFQKAKERLEAKHRMSQVMREWEAEARQA 420
QY 421 KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
Db 421 KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
QY 481 QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKAAQIRSQVWTHLRVIYER 540
Db 481 QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKAAQIRSQVWTHLRVIYER 540
QY 541 MNQSLLLYNVPAVAEIQDEVDLQKEQNSDDVLANNISEPRISYNDALMPSLTET 600
Db 541 MNQSLLLYNVPAVAEIQDEVDLQKEQNSDDVLANNISEPRISYNDALMPSLTET 600
QY 601 KTTVELLPVNGEFLDLDLQPHWSFGADSVPAANTENEVEPVDARPAADRGLTTRPGSGLTN 660
Db 601 KTTVELLPVNGEFLDLDLQPHWSFGADSVPAANTENEVEPVDARPAADRGLTTRPGSGLTN 660
QY 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNGKGAIIGLMVGWVIATVIVITL 720
Db 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNGKGAIIGLMVGWVIATVIVITL 720
QY 721 VMLKKQYTSIIHGVVEVDAAVTPEERHLSKMQONGYENPTYKFEQMQN 770
Db 721 VMLKKQYTSIIHGVVEVDAAVTPEERHLSKMQONGYENPTYKFEQMQN 770

RESULT 10
US-09-794-927-59
; Sequence 59, Application US/09794927
; Patent No. US20010016324A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Heinrichson, Robert L.
; APPLICANT: Parodi, Luis A.
; APPLICANT: Yan, Riqiang
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
; TITLE OF INVENTION: USES
; FILE REFERENCE: 28341/6280FG
; CURRENT APPLICATION NUMBER: US/09/794, 927
; CURRENT FILING DATE: 2001-02-27
; PRIOR FILING DATE: 1999-10-13
; PRIOR FILING DATE: 1999-10-13
; PRIOR FILING DATE: 1999-09-23
; PRIOR FILING DATE: 1999-09-23
; PRIOR FILING DATE: 1999-09-23
; PRIOR FILING DATE: 1999-09-23
; PRIOR FILING DATE: 1999-09-23
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 59
; LENGTH: 772
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-794-927-59

Query Match 100.0%; Score 4058; DB 10; Length 772;
Best Local Similarity 100.0%; Pred. No. 3.9e-267;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLPGLALLLAANTARALEVPTDGNAGLLAEQIAIFCCRLNMHNVQNGKWDSPSGTK 60
Db 1 MLPGLALLLAANTARALEVPTDGNAGLLAEQIAIFCCRLNMHNVQNGKWDSPSGTK 60
```

```
QY 61 TCIDTKEGILQYCOBYPELOITNVVNEANQVPTTQNMCKRGRKCKKTHPHFVIRCLVG 120
Db 61 TCIDTKEGILQYCOBYPELOITNVVNEANQVPTTQNMCKRGRKCKKTHPHFVIRCLVG 120
QY 121 EFVSDALLVPCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPGGIDKFR 180
Db 121 EFVSDALLVPCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPGGIDKFR 180
QY 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEVEEAEVEE 240
Db 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEVEEAEVEE 240
QY 241 EADDDDEDDGDEVEEAEPEYEATERTTSIATTTTTTSTESVEEVREVCSQAETGPC 300
Db 241 EADDDDEDDGDEVEEAEPEYEATERTTSIATTTTTTSTESVEEVREVCSQAETGPC 300
QY 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNFDTEYCMVCGSAMQSLLKTTQEPPLARD 360
Db 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNFDTEYCMVCGSAMQSLLKTTQEPPLARD 360
QY 361 PVKLPPTAASTPDVADKYLETGPDENEHAHFQKAKERLEAKHRMSQVMREWEAEARQA 420
Db 361 PVKLPPTAASTPDVADKYLETGPDENEHAHFQKAKERLEAKHRMSQVMREWEAEARQA 420
QY 421 KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
Db 421 KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
QY 481 QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKAAQIRSQVWTHLRVIYER 540
Db 481 QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKAAQIRSQVWTHLRVIYER 540
QY 541 MNQSLLLYNVPAVAEIQDEVDLQKEQNSDDVLANNISEPRISYNDALMPSLTET 600
Db 541 MNQSLLLYNVPAVAEIQDEVDLQKEQNSDDVLANNISEPRISYNDALMPSLTET 600
QY 601 KTTVELLPVNGEFLDLDLQPHWSFGADSVPAANTENEVEPVDARPAADRGLTTRPGSGLTN 660
Db 601 KTTVELLPVNGEFLDLDLQPHWSFGADSVPAANTENEVEPVDARPAADRGLTTRPGSGLTN 660
QY 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNGKGAIIGLMVGWVIATVIVITL 720
Db 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNGKGAIIGLMVGWVIATVIVITL 720
QY 721 VMLKKQYTSIIHGVVEVDAAVTPEERHLSKMQONGYENPTYKFEQMQN 770
Db 721 VMLKKQYTSIIHGVVEVDAAVTPEERHLSKMQONGYENPTYKFEQMQN 770

RESULT 11
US-09-795-847-59
; Sequence 59, Application US/09795847
; Patent No. US20010018208A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Heinrichson, Robert L.
; APPLICANT: Parodi, Luis A.
; APPLICANT: Yan, Riqiang
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
; TITLE OF INVENTION: USES
; FILE REFERENCE: 28341/6280DE
; CURRENT APPLICATION NUMBER: US/09/795, 847
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 09/416, 901
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: 60/155, 493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/404, 133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
```

; PRIOR APPLICATION NUMBER: 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 59
; LENGTH: 772
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-795-847-59

Query Match 100.0%; Score 4058; DB 10; Length 772;
Best Local Similarity 100.0%; Pred. No. 3.9e-267;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MLPGLALLLAATAWTAARALEVPTDGNAGLLAEPQIAMFCGRNLNHNHNVQNGKWDSPSGTK 60
Db 1 MLPGLALLLAATAWTAARALEVPTDGNAGLLAEPQIAMFCGRNLNHNHNVQNGKWDSPSGTK 60
QY 61 TCIDTREGILQYCOEYVPELQITNVVEANQPVTIQNWKGRKQCKTHPHFVYPRCLVG 120
Db 61 TCIDTREGILQYCOEYVPELQITNVVEANQPVTIQNWKGRKQCKTHPHFVYPRCLVG 120
QY 121 EFVSDALLVPDKCKFLHQERMDVCETHLHHVTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
Db 121 EFVSDALLVPDKCKFLHQERMDVCETHLHHVTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
QY 181 GVEFVCCPLAESDNVDSADAEEDSDVWGGADTDYADGSEDKVVEVAEEVEAEVEE 240
Db 181 GVEFVCCPLAESDNVDSADAEEDSDVWGGADTDYADGSEDKVVEVAEEVEAEVEE 240
QY 241 EADDDDEDDGDEVEEAEPEYEATERTTSIATTTTTTESVEEVVRCVSEQAETGPC 300
Db 241 EADDDDEDDGDEVEEAEPEYEATERTTSIATTTTTTESVEEVVRCVSEQAETGPC 300
QY 301 RAMISRWFYDVTGKCAPFFYGGCGGNRNNDTEECVMVCGSAMSQSLLKTTQEPPLARD 360
Db 301 RAMISRWFYDVTGKCAPFFYGGCGGNRNNDTEECVMVCGSAMSQSLLKTTQEPPLARD 360
QY 361 PVKLPPTAASTPDVADVKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAEERQA 420
Db 361 PVKLPPTAASTPDVADVKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAEERQA 420
QY 421 KNLPAKADKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
Db 421 KNLPAKADKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
QY 481 QAVPPRPRHVENMLKKYVRAEQKDRQHTLKHFEHVRVDPKKAQIRSQVWTHLRVIYER 540
Db 481 QAVPPRPRHVENMLKKYVRAEQKDRQHTLKHFEHVRVDPKKAQIRSQVWTHLRVIYER 540
QY 541 MNQSLLLYNVPAVAEEIQDEVDLQEKQNYSDVLANMISEPRIISYNDALMPSLTET 600
Db 541 MNQSLLLYNVPAVAEEIQDEVDLQEKQNYSDVLANMISEPRIISYNDALMPSLTET 600
QY 601 KTTVELLPVNGEFLDQLPWHISFGADSVPAANTEVEVPDARPAADRLGLTRPGSGLTN 660
Db 601 KTTVELLPVNGEFLDQLPWHISFGADSVPAANTEVEVPDARPAADRLGLTRPGSGLTN 660
QY 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFFAEDVGSNKGALIGLMVGWGVVATVIVITL 720
Db 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFFAEDVGSNKGALIGLMVGWGVVATVIVITL 720
QY 721 VMLKKQYTSIIHGVVEVDAAVTPPEERHLSKMQQNGYENPTYKFFEQMON 770
Db 721 VMLKKQYTSIIHGVVEVDAAVTPPEERHLSKMQQNGYENPTYKFFEQMON 770
```

RESULT 12

US-09-794-743-59
; Sequence 59, Application US/09794743
; Patent No. US20010021391A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.

; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Heinrikson, Robert L.
; APPLICANT: Parodi, Luis A.
; APPLICANT: Yan, Riqiang
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
; TITLE OF INVENTION: USES
; FILE REFERENCE: 28341/6280BC
; CURRENT APPLICATION NUMBER: US/09/794,743
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 09/416,901
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 59
; LENGTH: 772
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-794-743-59

Query Match 100.0%; Score 4058; DB 10; Length 772;
Best Local Similarity 100.0%; Pred. No. 3.9e-267;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MLPGLALLLAATAWTAARALEVPTDGNAGLLAEPQIAMFCGRNLNHNHNVQNGKWDSPSGTK 60
Db 1 MLPGLALLLAATAWTAARALEVPTDGNAGLLAEPQIAMFCGRNLNHNHNVQNGKWDSPSGTK 60
QY 61 TCIDTREGILQYCOEYVPELQITNVVEANQPVTIQNWKGRKQCKTHPHFVYPRCLVG 120
Db 61 TCIDTREGILQYCOEYVPELQITNVVEANQPVTIQNWKGRKQCKTHPHFVYPRCLVG 120
QY 121 EFVSDALLVPDKCKFLHQERMDVCETHLHHVTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
Db 121 EFVSDALLVPDKCKFLHQERMDVCETHLHHVTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
QY 181 GVEFVCCPLAESDNVDSADAEEDSDVWGGADTDYADGSEDKVVEVAEEVEAEVEE 240
Db 181 GVEFVCCPLAESDNVDSADAEEDSDVWGGADTDYADGSEDKVVEVAEEVEAEVEE 240
QY 241 EADDDDEDDGDEVEEAEPEYEATERTTSIATTTTTTESVEEVVRCVSEQAETGPC 300
Db 241 EADDDDEDDGDEVEEAEPEYEATERTTSIATTTTTTESVEEVVRCVSEQAETGPC 300
QY 301 RAMISRWFYDVTGKCAPFFYGGCGGNRNNDTEECVMVCGSAMSQSLLKTTQEPPLARD 360
Db 301 RAMISRWFYDVTGKCAPFFYGGCGGNRNNDTEECVMVCGSAMSQSLLKTTQEPPLARD 360
QY 361 PVKLPPTAASTPDVADVKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAEERQA 420
Db 361 PVKLPPTAASTPDVADVKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAEERQA 420
QY 421 KNLPAKADKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
Db 421 KNLPAKADKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
QY 481 QAVPPRPRHVENMLKKYVRAEQKDRQHTLKHFEHVRVDPKKAQIRSQVWTHLRVIYER 540
Db 481 QAVPPRPRHVENMLKKYVRAEQKDRQHTLKHFEHVRVDPKKAQIRSQVWTHLRVIYER 540
QY 541 MNQSLLLYNVPAVAEEIQDEVDLQEKQNYSDVLANMISEPRIISYNDALMPSLTET 600
Db 541 MNQSLLLYNVPAVAEEIQDEVDLQEKQNYSDVLANMISEPRIISYNDALMPSLTET 600
QY 601 KTTVELLPVNGEFLDQLPWHISFGADSVPAANTEVEVPDARPAADRLGLTRPGSGLTN 660
```

Db 601 KTTVELLPVNGEFLDLPQWHSFGADSVPAANTENEVEPVDARPAADRLTTRPGSGLTN 660
QY 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVGGVVIATVIVITL 720
Db 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVGGVVIATVIVITL 720
QY 721 VMLKKKQYTSIIHGVVEVDAVTPPEERHLSKMQONGYENPTYKFFEQMON 770
Db 721 VMLKKKQYTSIIHGVVEVDAVTPPEERHLSKMQONGYENPTYKFFEQMON 770

RESULT 13

US-09-794-748-59
; Sequence 59, Application US/09794748
; Patent No. US20020037315A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Heinrichson, Robert L.
; APPLICANT: Parodi, Luis A.
; APPLICANT: Yan, Riqiang
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
; TITLE OF INVENTION: USES
; FILE REFERENCE: 28341/6280JL
; CURRENT APPLICATION NUMBER: US/09/794,748
; CURRENT FILING DATE: 2001-02-27
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 59
; LENGTH: 772
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-794-748-59

Query Match 100.0%; Score 4058; DB 10; Length 772;
Best Local Similarity 100.0%; Pred. No. 3.9e-267;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLPGLALLLAAWTARALEVPTDGNAGLAEPOIAMFCGRLNHNHNVQNGKWDSPSGTK 60
Db 1 MLPGLALLLAAWTARALEVPTDGNAGLAEPOIAMFCGRLNHNHNVQNGKWDSPSGTK 60
QY 61 TCIDTKEGILQYCEVYPELQITNVVEANQPVTIQNCKRGRKQCKTHPHFVPIYRCLVG 120
Db 61 TCIDTKEGILQYCEVYPELQITNVVEANQPVTIQNCKRGRKQCKTHPHFVPIYRCLVG 120
QY 121 EFVSDALLVPDKCKFLHQRMDVCEVTHLHWTAKETCSKSNLHNDYGMLLPCGIDKFR 180
Db 121 EFVSDALLVPDKCKFLHQRMDVCEVTHLHWTAKETCSKSNLHNDYGMLLPCGIDKFR 180
QY 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSDKVVVEAEVEE 240
Db 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSDKVVVEAEVEE 240
QY 241 EADDEDEDGEVEEAEPEEATERTTSIATTTTTTTSVEEVYREVYCSQAETGPC 300
Db 241 EADDEDEDGEVEEAEPEEATERTTSIATTTTTTTSVEEVYREVYCSQAETGPC 300
QY 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNFDTEECYCMVCGSAMSOLLKTTQEPILARD 360
Db 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNFDTEECYCMVCGSAMSOLLKTTQEPILARD 360

QY 361 PVKLTPTTAASTPDVDKYLETGPDENEHAHFOKAKERLEAKHRRMSQVMREWEAEQOA 420
Db 361 PVKLTPTTAASTPDVDKYLETGPDENEHAHFOKAKERLEAKHRRMSQVMREWEAEQOA 420
QY 421 KNLPKADKAVIQHFQEKVESLEQEAANERQOLVETHMARVEAMLNDRRLALENYITAL 480
Db 421 KNLPKADKAVIQHFQEKVESLEQEAANERQOLVETHMARVEAMLNDRRLALENYITAL 480
QY 481 QAVPPRPRHVFNNLKKYVRAEQOKDRQHTLKHFEHVMVDPKKAQIRSQVTHLRVIYER 540
Db 481 QAVPPRPRHVFNNLKKYVRAEQOKDRQHTLKHFEHVMVDPKKAQIRSQVTHLRVIYER 540
QY 541 MNQSLSLYNNPVAEEIQDEVDLLOKEONYSDOVLANMISEPRISYGNDAIMPSTLET 600
Db 541 MNQSLSLYNNPVAEEIQDEVDLLOKEONYSDOVLANMISEPRISYGNDAIMPSTLET 600
QY 601 KTTVELLPVNGEFLDLPQWHSFGADSVPAANTENEVEPVDARPAADRLTTRPGSGLTN 660
Db 601 KTTVELLPVNGEFLDLPQWHSFGADSVPAANTENEVEPVDARPAADRLTTRPGSGLTN 660
QY 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVGGVVIATVIVITL 720
Db 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVGGVVIATVIVITL 720
QY 721 VMLKKKQYTSIIHGVVEVDAVTPPEERHLSKMQONGYENPTYKFFEQMON 770
Db 721 VMLKKKQYTSIIHGVVEVDAVTPPEERHLSKMQONGYENPTYKFFEQMON 770

RESULT 14

US-09-794-925-59
; Sequence 59, Application US/09794925
; Patent No. US20020064819A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Heinrichson, Robert L.
; APPLICANT: Parodi, Luis A.
; APPLICANT: Yan, Riqiang
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND U
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 28341/6280HI
; CURRENT APPLICATION NUMBER: US/09/794,925
; CURRENT FILING DATE: 2001-02-27
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 59
; LENGTH: 772
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-794-925-59

Query Match 100.0%; Score 4058; DB 10; Length 772;
Best Local Similarity 100.0%; Pred. No. 3.9e-267;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLPGLALLLAAWTARALEVPTDGNAGLAEPOIAMFCGRLNHNHNVQNGKWDSPSGTK 60
Db 1 MLPGLALLLAAWTARALEVPTDGNAGLAEPOIAMFCGRLNHNHNVQNGKWDSPSGTK 60
QY 61 TCIDTKEGILQYCEVYPELQITNVVEANQPVTIQNCKRGRKQCKTHPHFVPIYRCLVG 120
Db 61 TCIDTKEGILQYCEVYPELQITNVVEANQPVTIQNCKRGRKQCKTHPHFVPIYRCLVG 120

Db 61 TCIDTKEGILQVCQEVYPELQITNVVVEANQVPTIQNWCKRGKCKTHPHFVPIYRCLVG 120
Qy 121 EFVSDALLVPDKCFHQRMDVCEETHLHWHVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
Db 121 EFVSDALLVPDKCFHQRMDVCEETHLHWHVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
Qy 181 GVEFVCCPLAEBSDNVSDAEDSDVWVGADTDYADGSEDKVVEVAEEVAEVEE 240
Db 181 GVEFVCCPLAEBSDNVSDAEDSDVWVGADTDYADGSEDKVVEVAEEVAEVEE 240
Qy 241 EADDDDEDDGDEVEEAEEPEEATERTTSIATTTTTTTSVEEVREVSCQAETGPC 300
Db 241 EADDDDEDDGDEVEEAEEPEEATERTTSIATTTTTTTSVEEVREVSCQAETGPC 300
Qy 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNFDTTEYCMVCGSAMSLLKTTQEP LARD 360
Db 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNFDTTEYCMVCGSAMSLLKTTQEP LARD 360
Qy 361 PVKLPPTAASPDVADVKYLETTPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAEERQA 420
Db 361 PVKLPPTAASPDVADVKYLETTPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAEERQA 420
Qy 421 KNLPKADKAVIQHFQKVESLEQEAANERQOLVETHMARVEAMLNDRRRLALENYITAL 480
Db 421 KNLPKADKAVIQHFQKVESLEQEAANERQOLVETHMARVEAMLNDRRRLALENYITAL 480
Qy 481 QAVPPRPRHVFNMLKYYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 540
Db 481 QAVPPRPRHVFNMLKYYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 540
Qy 541 MNQSLSLLYNPVPAVEIEQDVEBLLQKEQNYSDVLANMISEPRISYGNDAIMPSTET 600
Db 541 MNQSLSLLYNPVPAVEIEQDVEBLLQKEQNYSDVLANMISEPRISYGNDAIMPSTET 600
Qy 601 KTTVELLPVNGEFSLLDLPWHSFGADSVDPANTENEVEPVDARPAADRGLTTRPGSGLTN 660
Db 601 KTTVELLPVNGEFSLLDLPWHSFGADSVDPANTENEVEPVDARPAADRGLTTRPGSGLTN 660
Qy 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGVIATVITL 720
Db 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGVIATVITL 720
Qy 721 VMLKKKQYTSIHGGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN 770
Db 721 VMLKKKQYTSIHGGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN 770

RESULT 15

US-09-681-442-59

; Sequence 59, Application US/09681442
; Patent No. US20020081634A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Heinrikson, Robert L.
; APPLICANT: Parodi, Luis A.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 28341/6280FC
; CURRENT APPLICATION NUMBER: US/09/681,442
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 09/416,901
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 59

; LENGTH: 772

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-681-442-59

Query Match

100.0%; Score 4058; DB 10; Length 772;

Best Local Similarity 100.0%; Pred. No. 3.9e-267;

Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPCIAMFCGRLNMHNMQKWDSPSGTK 60

Db 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPCIAMFCGRLNMHNMQKWDSPSGTK 60

Qy 61 TCIDTKEGILQVCQEVYPELQITNVVVEANQVPTIQNWCKRGKCKTHPHFVPIYRCLVG 120

Db 61 TCIDTKEGILQVCQEVYPELQITNVVVEANQVPTIQNWCKRGKCKTHPHFVPIYRCLVG 120

Qy 121 EFVSDALLVPDKCFHQRMDVCEETHLHWHVAKETCSEKSTNLHDYGMLLPCGIDKFR 180

Db 121 EFVSDALLVPDKCFHQRMDVCEETHLHWHVAKETCSEKSTNLHDYGMLLPCGIDKFR 180

Qy 181 GVEFVCCPLAEBSDNVSDAEDSDVWVGADTDYADGSEDKVVEVAEEVAEVEE 240

Db 181 GVEFVCCPLAEBSDNVSDAEDSDVWVGADTDYADGSEDKVVEVAEEVAEVEE 240

Qy 241 EADDDDEDDGDEVEEAEEPEEATERTTSIATTTTTTTSVEEVREVSCQAETGPC 300

Db 241 EADDDDEDDGDEVEEAEEPEEATERTTSIATTTTTTTSVEEVREVSCQAETGPC 300

Qy 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNFDTTEYCMVCGSAMSQSLLKTTQEP LARD 360

Db 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNFDTTEYCMVCGSAMSQSLLKTTQEP LARD 360

Qy 361 PVKLPPTAASPDVADVKYLETTPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAEERQA 420

Db 361 PVKLPPTAASPDVADVKYLETTPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAEERQA 420

Qy 421 KNLPKADKAVIQHFQKVESLEQEAANERQOLVETHMARVEAMLNDRRRLALENYITAL 480

Db 421 KNLPKADKAVIQHFQKVESLEQEAANERQOLVETHMARVEAMLNDRRRLALENYITAL 480

Qy 481 QAVPPRPRHVFNMLKYYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 540

Db 481 QAVPPRPRHVFNMLKYYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 540

Qy 541 MNQSLSLLYNPVPAVEIEQDVEBLLQKEQNYSDVLANMISEPRISYGNDAIMPSTET 600

Db 541 MNQSLSLLYNPVPAVEIEQDVEBLLQKEQNYSDVLANMISEPRISYGNDAIMPSTET 600

Qy 601 KTTVELLPVNGEFSLLDLPWHSFGADSVDPANTENEVEPVDARPAADRGLTTRPGSGLTN 660

Db 601 KTTVELLPVNGEFSLLDLPWHSFGADSVDPANTENEVEPVDARPAADRGLTTRPGSGLTN 660

Qy 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGVIATVITL 720

Db 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGVIATVITL 720

Qy 721 VMLKKKQYTSIHGGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN 770

Db 721 VMLKKKQYTSIHGGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN 770

Search completed: January 28, 2003, 14:40:06

Job time : 13.0956 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:10:57 ; Search time 15,9847 Seconds
(without alignments)
1417.333 Million cell updates/sec

Title: US-09-904-987-2

Perfect score: 4058

Sequence: 1 MLPGLALLLLAAWTARALEV.....RMQNGYENPTYKFEQMQN 770

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*

- 1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4058	100.0	770	1 US-08-133-248-8	Sequence 8, Appli
2	4058	100.0	770	1 US-08-231-940-1	Sequence 1, Appli
3	4058	100.0	770	2 US-08-641-774-1	Sequence 1, Appli
4	4058	100.0	770	2 US-08-104-165-3	Sequence 3, Appli
5	4058	100.0	770	3 US-08-464-250-3	Sequence 3, Appli
6	4058	100.0	770	4 US-08-464-250-3	Sequence 3, Appli
7	4058	100.0	770	4 US-09-548-372D-55	Sequence 55, Appli
8	4058	100.0	770	4 US-09-548-367D-55	Sequence 55, Appli
9	4058	100.0	772	4 US-09-548-372D-59	Sequence 59, Appli
10	4058	100.0	772	4 US-09-548-367D-59	Sequence 59, Appli
11	3943.5	97.2	751	1 US-08-123-702-4	Sequence 4, Appli
12	3943.5	97.2	751	2 US-08-104-165-2	Sequence 2, Appli
13	3943.5	97.2	751	2 US-08-422-333-2	Sequence 2, Appli
14	3943.5	97.2	751	2 US-08-422-333-21	Sequence 21, Appli
15	3943.5	97.2	751	3 US-08-464-250-2	Sequence 2, Appli
16	3943.5	97.2	751	4 US-08-464-250-2	Sequence 2, Appli
17	3943.5	97.2	751	4 US-08-832-867-5	Sequence 5, Appli
18	3943.5	97.2	751	4 US-09-548-372D-57	Sequence 57, Appli
19	3943.5	97.2	751	4 US-09-548-367D-57	Sequence 57, Appli
20	3943.5	97.2	751	6 5187153-2	Patent No. 5187153
21	3943.5	97.2	751	6 5223482-2	Patent No. 5223482
22	3943.5	97.2	753	4 US-09-548-372D-61	Sequence 61, Appli
23	3943.5	97.2	753	4 US-09-548-367D-61	Sequence 61, Appli
24	3937.5	97.0	751	6 5220013-2	Patent No. 5220013
25	3590.5	88.5	695	1 US-08-123-702-2	Sequence 2, Appli
26	3590.5	88.5	695	2 US-08-104-165-1	Sequence 1, Appli
27	3590.5	88.5	695	3 US-08-464-250-1	Sequence 1, Appli

28 3590.5 88.5 695 4 US-08-464-250-1 Sequence 1, Appli
29 3590.5 88.5 695 4 US-09-458-481B-7 Sequence 7, Appli
30 3590.5 88.5 695 4 US-09-458-481B-8 Sequence 8, Appli
31 3590.5 88.5 695 4 US-09-548-372D-10 Sequence 10, Appli
32 3590.5 88.5 695 4 US-09-548-367D-10 Sequence 10, Appli
33 3590.5 88.5 695 6 5218100-2 Patent No. 5218100
34 3590.5 88.5 697 4 US-09-548-372D-16 Sequence 16, Appli
35 3590.5 88.5 697 4 US-09-548-367D-16 Sequence 16, Appli
36 3585.5 88.4 695 4 US-09-548-372D-14 Sequence 14, Appli
37 3585.5 88.4 695 4 US-09-548-367D-14 Sequence 14, Appli
38 3585.5 88.4 697 4 US-09-548-372D-20 Sequence 20, Appli
39 3585.5 88.4 697 4 US-09-548-367D-20 Sequence 20, Appli
40 3584.5 88.3 694 1 US-08-339-152A-18 Sequence 18, Appli
41 3584.5 88.3 694 2 US-08-007-999B-5 Sequence 5, Appli
42 3584.5 88.3 694 2 US-08-689-276A-5 Sequence 5, Appli
43 3582.5 88.3 695 4 US-09-548-372D-12 Sequence 12, Appli
44 3582.5 88.3 695 4 US-09-548-367D-12 Sequence 12, Appli
45 3582.5 88.3 697 4 US-09-548-372D-18 Sequence 18, Appli

ALIGNMENTS

RESULT 1

US-08-133-248-8

; Sequence 8, Application US/08133248

; Patent No. 5525714

; GENERAL INFORMATION:

; APPLICANT:

; TITLE OF INVENTION: MUTATED FORM OF THE BETA-AMYLOID PRECURSOR

; NUMBER OF SEQUENCES: 8

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/133,248

; INFORMATION FOR SEQ ID NO: 8:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 770 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-133-248-8

Query Match 100.0%; Score 4058; DB 1; Length 770;
Best Local Similarity 100.0%; Pred. No. 1.1e-280;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPTAMFCGRLLNMNMVONGKWDSPSGTK 60

Db 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPTAMFCGRLLNMNMVONGKWDSPSGTK 60

QY 61 TCIDTKEGILQYQCBVPELQITNVVEANQPVITQNWCKRGKCKTTPHVPVPCVLVG 120

Db 61 TCIDTKEGILQYQCBVPELQITNVVEANQPVITQNWCKRGKCKTTPHVPVPCVLVG 120

QY 121 EFVSDALLVPDKCFKHQGRMDVCEVTHLHWHTVAKETSEKSTNLHDYGMLLPCGIDKFR 180

Db 121 EFVSDALLVPDKCFKHQGRMDVCEVTHLHWHTVAKETSEKSTNLHDYGMLLPCGIDKFR 180

QY 181 GVEFFCCPLAESDNVDSADAEEDSDVWVGADTDYADGEDKVVVEVAEEVAEEVEE 240

Db 181 GVEFFCCPLAESDNVDSADAEEDSDVWVGADTDYADGEDKVVVEVAEEVAEEVEE 240

QY 241 EADDDDEDDGDEVEEAEEPEEATERTTSTATTTTTTTSVEEVREVCSQAETGPC 300

Db 241 EADDDDEDDGDEVEEAEEPEEATERTTSTATTTTTTTSVEEVREVCSQAETGPC 300

QY 301 RAMISRWFYDVTGKCAPFFYGGCGGNRNFDTEECYMAVCGSAMQSLLKTTQEP LARD 360

Db 301 RAMISRWFYDVTGKCAPFFYGGCGGNRNFDTEECYMAVCGSAMQSLLKTTQEP LARD 360

```
Db 301 RAMISRWFVDTGKCAPFYGGGNNRNFDEEYCMVCGSAMQSLLKTTQEPPLARD 360
Qy 361 PVKLPTTAASPTDAVDKYLETDPDENEHAHFQKAKERLEAKHRMSQVMREWEAEARQA 420
Db 361 PVKLPTTAASPTDAVDKYLETDPDENEHAHFQKAKERLEAKHRMSQVMREWEAEARQA 420
Qy 421 KNLPKADKKAIVQHFKQKVESLQEAANERQQLVETHMARVEAMLNDRRLALENYITL 480
Db 421 KNLPKADKKAIVQHFKQKVESLQEAANERQQLVETHMARVEAMLNDRRLALENYITL 480
Qy 481 QAVPPRPRHVFNMLKKYVRAEQDRQHTLKHFEHVRMDPKKAAQIRSQVMTHLRVIYER 540
Db 481 QAVPPRPRHVFNMLKKYVRAEQDRQHTLKHFEHVRMDPKKAAQIRSQVMTHLRVIYER 540
Qy 541 MNQSLSLLYNVPAAVEIQDEVDLQKEQNSDDVLANNMISEPRISYGNDAIMPSTET 600
Db 541 MNQSLSLLYNVPAAVEIQDEVDLQKEQNSDDVLANNMISEPRISYGNDAIMPSTET 600
Qy 601 KTTVELLPVNGEFSLDDLPWHSFGADSVDPANTENEVEPVDARPAADRGLTTRPGSGLTN 660
Db 601 KTTVELLPVNGEFSLDDLPWHSFGADSVDPANTENEVEPVDARPAADRGLTTRPGSGLTN 660
Qy 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVGWVIATVIVITL 720
Db 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVGWVIATVIVITL 720
Qy 721 VMLKKKQYTSIHGGVVEVDAVTPPEERHLSKMQONGYENPTYKFFEQMON 770
Db 721 VMLKKKQYTSIHGGVVEVDAVTPPEERHLSKMQONGYENPTYKFFEQMON 770

RESULT 2
US-08-231-940-1
; Sequence 1, Application US/08231940
; Patent No. 5550216
; GENERAL INFORMATION:
; APPLICANT: MIYAZAKI, Kaoru
; TITLE OF INVENTION: GELATINASE A INHIBITOR
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/231,940
; FILING DATE: 25-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 5-120457
; FILING DATE: 26-APR-1993
; PRIOR APPLICATION NUMBER: JP 6-62129
; FILING DATE: 08-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: NEIMARK, Sheridan
; REGISTRATION NUMBER: 20,520
; REFERENCE/DOCKET NUMBER: MIYAZAKI-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 770 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
```

```
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-231-940-1

Query Match 100.0%; Score 4058; DB 1; Length 770;
Best Local Similarity 100.0%; Pred. No. 1.1e-280;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLPLGLALLLAATAWARALEVPTDGNAGLLAEPQIATMFCGRNLNMHNWONGKWDSDPSGK 60
Db 1 MLPLGLALLLAATAWARALEVPTDGNAGLLAEPQIATMFCGRNLNMHNWONGKWDSDPSGK 60
Qy 61 TCIDTKEGILQYCOBVPPELQITNNVEANQPVITQNMCKRGKCKCTKHPHEVIPRCLVG 120
Db 61 TCIDTKEGILQYCOBVPPELQITNNVEANQPVITQNMCKRGKCKCTKHPHEVIPRCLVG 120
Qy 121 EFVSALLVPKCKFLHQRMDVCEVTHLHWTVAKETSEKSTNLHDYGMLLPGCIDKFR 180
Db 121 EFVSALLVPKCKFLHQRMDVCEVTHLHWTVAKETSEKSTNLHDYGMLLPGCIDKFR 180
Qy 181 GVEFYCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEEVEE 240
Db 181 GVEFYCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEEVEE 240
Qy 241 EADDEDEDEDEVEEAEPEEATERTTSIATTTTTTSTESVEEVREVCSQAETGPC 300
Db 241 EADDEDEDEDEVEEAEPEEATERTTSIATTTTTTSTESVEEVREVCSQAETGPC 300
Qy 301 RAMISRWFVDTGKCAPFYGGGNNRNFDEEYCMVCGSAMQSLLKTTQEPPLARD 360
Db 301 RAMISRWFVDTGKCAPFYGGGNNRNFDEEYCMVCGSAMQSLLKTTQEPPLARD 360
Qy 361 PVKLPTTAASPTDAVDKYLETDPDENEHAHFQKAKERLEAKHRMSQVMREWEAEARQA 420
Db 361 PVKLPTTAASPTDAVDKYLETDPDENEHAHFQKAKERLEAKHRMSQVMREWEAEARQA 420
Qy 421 KNLPKADKKAIVQHFKQKVESLQEAANERQQLVETHMARVEAMLNDRRLALENYITL 480
Db 421 KNLPKADKKAIVQHFKQKVESLQEAANERQQLVETHMARVEAMLNDRRLALENYITL 480
Qy 481 QAVPPRPRHVFNMLKKYVRAEQDRQHTLKHFEHVRMDPKKAAQIRSQVMTHLRVIYER 540
Db 481 QAVPPRPRHVFNMLKKYVRAEQDRQHTLKHFEHVRMDPKKAAQIRSQVMTHLRVIYER 540
Qy 541 MNQSLSLLYNVPAAVEIQDEVDLQKEQNSDDVLANNMISEPRISYGNDAIMPSTET 600
Db 541 MNQSLSLLYNVPAAVEIQDEVDLQKEQNSDDVLANNMISEPRISYGNDAIMPSTET 600
Qy 601 KTTVELLPVNGEFSLDDLPWHSFGADSVDPANTENEVEPVDARPAADRGLTTRPGSGLTN 660
Db 601 KTTVELLPVNGEFSLDDLPWHSFGADSVDPANTENEVEPVDARPAADRGLTTRPGSGLTN 660
Qy 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVGWVIATVIVITL 720
Db 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVGWVIATVIVITL 720
Qy 721 VMLKKKQYTSIHGGVVEVDAVTPPEERHLSKMQONGYENPTYKFFEQMON 770
Db 721 VMLKKKQYTSIHGGVVEVDAVTPPEERHLSKMQONGYENPTYKFFEQMON 770

RESULT 3
US-08-641-774-1
; Sequence 1, Application US/08641774
; Patent No. 5843695
; GENERAL INFORMATION:
; APPLICANT: MIYAZAKI, Kaoru
; TITLE OF INVENTION: BETA-AP DECOMPOSING AGENT
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
```

STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/641,774
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/232,474
FILING DATE: 25-APR-1994
APPLICATION NUMBER: JP 5-122207
FILING DATE: 26-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-51133
FILING DATE: 08-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: NEIMARK, Sheridan
REGISTRATION NUMBER: 20,520
REFERENCE/DOCKET NUMBER: MIYAZAKI-3
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 770 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-641-774-1

Query Match 100.0%; Score 4058; DB 2; Length 770;
Best Local Similarity 100.0%; Pred. No. 1.le-280;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLPGLALLLAATAWATALEVPDGNAGLAEPQIAMFCGRNLNHNMVQNGKWDSPSGTK 60
DB 1 MLPGLALLLAATAWATALEVPDGNAGLAEPQIAMFCGRNLNHNMVQNGKWDSPSGTK 60

QY 61 TCIDTKEGILQYCOEYVPELQIINNVEANQPVTIQNCKRGRKQCKTHPHFVPIRYCLVG 120
DB 61 TCIDTKEGILQYCOEYVPELQIINNVEANQPVTIQNCKRGRKQCKTHPHFVPIRYCLVG 120

QY 121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETSEKSTNLHDYGMLLPCGIDKFR 180
DB 121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETSEKSTNLHDYGMLLPCGIDKFR 180

QY 181 GVEFVCCPLAESDNVDSADAEDDSVVMWGGADTDYADGSEDKVVEVAEEVEEVEE 240
DB 181 GVEFVCCPLAESDNVDSADAEDDSVVMWGGADTDYADGSEDKVVEVAEEVEEVEE 240

QY 241 EADDEDEDDGVEEAEPEEATERTSIATTTTTTSTESVEEVVREVCEQATGTC 300
DB 241 EADDEDEDDGVEEAEPEEATERTSIATTTTTTSTESVEEVVREVCEQATGTC 300

QY 301 RAMISRWYFDVTGKCAPFYGGCGNRNFDPEEYCMVCGSAMQSLLKTTQEBPLARD 360
DB 301 RAMISRWYFDVTGKCAPFYGGCGNRNFDPEEYCMVCGSAMQSLLKTTQEBPLARD 360

QY 361 PVKLPTTAASTPDVADKYLETGPDENEHAHFQKAKERLEAKHRMSQVMREWEAEERQA 420
DB 361 PVKLPTTAASTPDVADKYLETGPDENEHAHFQKAKERLEAKHRMSQVMREWEAEERQA 420

QY 421 KNLPKADKKAIVQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENITVAL 480
DB 421 KNLPKADKKAIVQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENITVAL 480

QY 481 QAVPPRPRHVNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 540

DB 481 QAVPPRPRHVNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 540
QY 541 MNQSLILYNPVAEEIQDEVDLLOKEQYSDDVLANMISEPRISYGNALMPSLTET 600
DB 541 MNQSLILYNPVAEEIQDEVDLLOKEQYSDDVLANMISEPRISYGNALMPSLTET 600
QY 601 KTTVELLPVNGEFSLDDLPWHSFGADSVPANTENEVEPVDARPAADRGLTRPGSGLTN 660
DB 601 KTTVELLPVNGEFSLDDLPWHSFGADSVPANTENEVEPVDARPAADRGLTRPGSGLTN 660
QY 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNGKGAIGLMVGGVVIATVITVL 720
DB 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNGKGAIGLMVGGVVIATVITVL 720
QY 721 VMLKKQYTSIHGVEVVEVDAVTPPEERHLSKMQONGYENPTYKFFEQMON 770
DB 721 VMLKKQYTSIHGVEVVEVDAVTPPEERHLSKMQONGYENPTYKFFEQMON 770

RESULT 4
US-08-104-165-3
; Sequence 3, Application US/08104165
; Patent No. 5877015
; GENERAL INFORMATION:
; APPLICANT: HARDY, John Anthony
; APPLICANT: GOATE, Alison Mary
; APPLICANT: MULLAN, Michael John
; APPLICANT: CHARTIER-HARLIN, Marie-Christine
; APPLICANT: OWEN, Michael John
; TITLE OF INVENTION: Test and Model for Alzheimer's Disease
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourle and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/104,165
; FILING DATE: 21-JAN-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 9101307.8
; FILING DATE: 21-JAN-1991
; APPLICATION NUMBER: 9118445.7
; FILING DATE: 28-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Liebeschuetz, Joe
; REGISTRATION NUMBER: 37,505
; REFERENCE/DOCKET NUMBER: 16163-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 770 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-104-165-3

Query Match 100.0%; Score 4058; DB 2; Length 770;
Best Local Similarity 100.0%; Pred. No. 1.le-280;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLPGLALLLAATAWATALEVPDGNAGLAEPQIAMFCGRNLNHNMVQNGKWDSPSGTK 60

QY 601 KTTVELLPVNGEFLDLDLPWHSGADSVDPANTENEVEPVDAADPGLTTRPGSGLTN 660
Db 601 KTTVELLPVNGEFLDLDLPWHSGADSVDPANTENEVEPVDAADPGLTTRPGSGLTN 660
QY 661 IKTEEISEVKNDAEFRHDSGYEVHHQKLVFFAEDVGSNGKAIIGLMVGGVVIATVITL 720
Db 661 IKTEEISEVKNDAEFRHDSGYEVHHQKLVFFAEDVGSNGKAIIGLMVGGVVIATVITL 720
QY 721 VMLKKKQYTSIHGGVVEVDAVTPPEERHLSKMQONGYENPTYKFFEQMON 770
Db 721 VMLKKKQYTSIHGGVVEVDAVTPPEERHLSKMQONGYENPTYKFFEQMON 770
RESULT 6
US-08-464-250-3
; Sequence 3, Application US/08464250
; Patent No. 6300540
; GENERAL INFORMATION:
; APPLICANT: HARDY, John Anthony
; GOATE, Allison Mary
; MULLAN, Michael John
; CHARTIER-HARLIN, Marie-Christine
; OWEN, Michael John
; TITLE OF INVENTION: Test and Model for Alzheimer's Disease
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/464,250
; FILING DATE: 05-Jun-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/104,165
; FILING DATE: 21-JAN-1992
; APPLICATION NUMBER: 9101307.8
; FILING DATE: 21-JAN-1991
; APPLICATION NUMBER: 9118445.7
; FILING DATE: 28-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Liebeschuetz, Joe
; REGISTRATION NUMBER: 37,505
; REFERENCE/DOCKET NUMBER: 16163-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 770 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-08-464-250-3
Query Match 100.0%; Score 4058; DB 4; Length 770;
Best Local Similarity 100.0%; Pred. No. 1.1e-280;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MLPGALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNHNMVQNGKWDSPSGTK 60
Db 1 MLPGALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNHNMVQNGKWDSPSGTK 60
QY 61 TCIDTKEGILQYCEVPELQITNVVEANQPVTIQNWKGRGRKQCKTHPHFVPIYRCLVG 120

Db 61 TCIDTKEGILQYCEVPELQITNVVEANQPVTIQNWKGRGRKQCKTHPHFVPIYRCLVG 120
QY 121 EFVSDALLVPDKCKFLHQRMDVCEYLHWHHTVAKETCSEKSTNLHDYGMLLPCGIDKPR 180
Db 121 EFVSDALLVPDKCKFLHQRMDVCEYLHWHHTVAKETCSEKSTNLHDYGMLLPCGIDKPR 180
QY 181 GYEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSDKVVVEAAEEVEAEVEE 240
Db 181 GYEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSDKVVVEAAEEVEAEVEE 240
QY 241 EADDDDEDDGDEVEEEAEPEYEATERTTSIATTTTTTTSVEEVEVVCSEQAETGPC 300
Db 241 EADDDDEDDGDEVEEEAEPEYEATERTTSIATTTTTTTSVEEVEVVCSEQAETGPC 300
QY 301 RAMISRWFYDVTGKCAPFFYGGCGGNRNFTTEYCHAVCSAMSQSLLKTTQEPPLARD 360
Db 301 RAMISRWFYDVTGKCAPFFYGGCGGNRNFTTEYCHAVCSAMSQSLLKTTQEPPLARD 360
QY 361 PVKLPPTAASTPDADVKKYLETPGDENEHAHFOKAKERLEAKHRERMSQVMEWEAEARQA 420
Db 361 PVKLPPTAASTPDADVKKYLETPGDENEHAHFOKAKERLEAKHRERMSQVMEWEAEARQA 420
QY 421 KNLPRADKXAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
Db 421 KNLPRADKXAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
QY 481 QAVPPRRHVFNNMLKKYVRAEQKDRQHTLKHFEHVMVMDPKKAAQIRSOVMTHLRYVER 540
Db 481 QAVPPRRHVFNNMLKKYVRAEQKDRQHTLKHFEHVMVMDPKKAAQIRSOVMTHLRYVER 540
QY 541 MNQSLSLYNPVAAVEIODEVDELLOKEQNYSDOVLANMISEPRISYNDALMPSLTET 600
Db 541 MNQSLSLYNPVAAVEIODEVDELLOKEQNYSDOVLANMISEPRISYNDALMPSLTET 600
QY 601 KTTVELLPVNGEFLDLDLPWHSGADSVDPANTENEVEPVDAADPGLTTRPGSGLTN 660
Db 601 KTTVELLPVNGEFLDLDLPWHSGADSVDPANTENEVEPVDAADPGLTTRPGSGLTN 660
QY 661 IKTEEISEVKNDAEFRHDSGYEVHHQKLVFFAEDVGSNGKAIIGLMVGGVVIATVITL 720
Db 661 IKTEEISEVKNDAEFRHDSGYEVHHQKLVFFAEDVGSNGKAIIGLMVGGVVIATVITL 720
QY 721 VMLKKKQYTSIHGGVVEVDAVTPPEERHLSKMQONGYENPTYKFFEQMON 770
Db 721 VMLKKKQYTSIHGGVVEVDAVTPPEERHLSKMQONGYENPTYKFFEQMON 770

RESULT 7

US-09-548-372D-55
; Sequence 55, Application US/09548372D
; Patent No. 6420534
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 29915/62801
; CURRENT APPLICATION NUMBER: US/09/548,372D
; CURRENT FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 55
; LENGTH: 770
; TYPE: PRT
; ORGANISM: Homo sapiens

US-09-548-372D-55

Query Match 100.0%; Score 4058; DB 4; Length 770;
Best Local Similarity 100.0%; Pred. No. 1.1e-280;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLPGLALLLAANTARALEVPTDGNAGLLAEPQIAMFCGRLLNMHNVQNGKWDSPSGTK 60
DB 1 MLPGLALLLAANTARALEVPTDGNAGLLAEPQIAMFCGRLLNMHNVQNGKWDSPSGTK 60
QY 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTIQNWKGRKQCKTHPHFVIPYRCLVG 120
DB 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTIQNWKGRKQCKTHPHFVIPYRCLVG 120
QY 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
DB 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
QY 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEPEEVEE 240
DB 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEPEEVEE 240
QY 241 EADDDDEDDGDEVEEAEPEEATERTTSIATTTTTTSTESVEEVREVCSQAETGPC 300
DB 241 EADDDDEDDGDEVEEAEPEEATERTTSIATTTTTTSTESVEEVREVCSQAETGPC 300
QY 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNFDEEYCMVCGSAMQSLLKTTOEPLARD 360
DB 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNFDEEYCMVCGSAMQSLLKTTOEPLARD 360
QY 361 PVKLPTTAASTPDAVDKYLETDPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAE 420
DB 361 PVKLPTTAASTPDAVDKYLETDPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAE 420
QY 421 KNLKADKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
DB 421 KNLKADKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
QY 481 QAVPPRPRHVNMLKKYVRAEQDKRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIER 540
DB 481 QAVPPRPRHVNMLKKYVRAEQDKRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIER 540
QY 541 MNQSLSLLYNPVAAEIQDEVELLQEQNSDDVLANNMISEPRISYGNDAIMPSTET 600
DB 541 MNQSLSLLYNPVAAEIQDEVELLQEQNSDDVLANNMISEPRISYGNDAIMPSTET 600
QY 601 KTTVELLPVNGEFSLDDLPWHSGADSVPAANTENEVEPVDARPAADRGLTTRPGSGLTN 660
DB 601 KTTVELLPVNGEFSLDDLPWHSGADSVPAANTENEVEPVDARPAADRGLTTRPGSGLTN 660
QY 661 IKTEEISEVMDAEFRHDSGYEVHHQKLVFAEDVGSNKGAIIGLMVGGVVIATVIVITL 720
DB 661 IKTEEISEVMDAEFRHDSGYEVHHQKLVFAEDVGSNKGAIIGLMVGGVVIATVIVITL 720
QY 721 VMLKKKQYTSIHGGVVEVDAVTPPEERHLSKMQONGYENPTYKFFEQMQN 770
DB 721 VMLKKKQYTSIHGGVVEVDAVTPPEERHLSKMQONGYENPTYKFFEQMQN 770

RESULT 8

US-09-548-367D-55
; Sequence 55, Application US/09548367D
; Patent No. 6440698
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 29915/6280H
; CURRENT APPLICATION NUMBER: US/09/548,367D
; CURRENT FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133

; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 55
; LENGTH: 770
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-548-367D-55

Query Match 100.0%; Score 4058; DB 4; Length 770;

Best Local Similarity 100.0%; Pred. No. 1.1e-280;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLPGLALLLAANTARALEVPTDGNAGLLAEPQIAMFCGRLLNMHNVQNGKWDSPSGTK 60
DB 1 MLPGLALLLAANTARALEVPTDGNAGLLAEPQIAMFCGRLLNMHNVQNGKWDSPSGTK 60
QY 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTIQNWKGRKQCKTHPHFVIPYRCLVG 120
DB 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTIQNWKGRKQCKTHPHFVIPYRCLVG 120
QY 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
DB 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
QY 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEPEEVEE 240
DB 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEPEEVEE 240
QY 241 EADDDDEDDGDEVEEAEPEEATERTTSIATTTTTTSTESVEEVREVCSQAETGPC 300
DB 241 EADDDDEDDGDEVEEAEPEEATERTTSIATTTTTTSTESVEEVREVCSQAETGPC 300
QY 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNFDEEYCMVCGSAMQSLLKTTOEPLARD 360
DB 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNFDEEYCMVCGSAMQSLLKTTOEPLARD 360
QY 361 PVKLPTTAASTPDAVDKYLETDPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAE 420
DB 361 PVKLPTTAASTPDAVDKYLETDPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAE 420
QY 421 KNLKADKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
DB 421 KNLKADKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
QY 481 QAVPPRPRHVNMLKKYVRAEQDKRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIER 540
DB 481 QAVPPRPRHVNMLKKYVRAEQDKRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIER 540
QY 541 MNQSLSLLYNPVAAEIQDEVELLQEQNSDDVLANNMISEPRISYGNDAIMPSTET 600
DB 541 MNQSLSLLYNPVAAEIQDEVELLQEQNSDDVLANNMISEPRISYGNDAIMPSTET 600
QY 601 KTTVELLPVNGEFSLDDLPWHSGADSVPAANTENEVEPVDARPAADRGLTTRPGSGLTN 660
DB 601 KTTVELLPVNGEFSLDDLPWHSGADSVPAANTENEVEPVDARPAADRGLTTRPGSGLTN 660
QY 661 IKTEEISEVMDAEFRHDSGYEVHHQKLVFAEDVGSNKGAIIGLMVGGVVIATVIVITL 720
DB 661 IKTEEISEVMDAEFRHDSGYEVHHQKLVFAEDVGSNKGAIIGLMVGGVVIATVIVITL 720
QY 721 VMLKKKQYTSIHGGVVEVDAVTPPEERHLSKMQONGYENPTYKFFEQMQN 770
DB 721 VMLKKKQYTSIHGGVVEVDAVTPPEERHLSKMQONGYENPTYKFFEQMQN 770

RESULT 9

US-09-548-372D-59
; Sequence 59, Application US/09548372D

Patent No. 6420534
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 29915/62801
; CURRENT APPLICATION NUMBER: US/09/548,372D
; CURRENT FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 59
; LENGTH: 772
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-548-372D-59

Query Match 100.0%; Score 4058; DB 4; Length 772;
Best Local Similarity 100.0%; Pred. No. 1.1e-280;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLPGLALLLAANTARALEVPTDGNAGLLAEPOIAMFCGRNLNMHNVQNGKWDSPSGTK 60
Db 1 MLPGLALLLAANTARALEVPTDGNAGLLAEPOIAMFCGRNLNMHNVQNGKWDSPSGTK 60

Qy 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTIQNCKRGRKCKOCTHPHFVPIYRCLVG 120
Db 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTIQNCKRGRKCKOCTHPHFVPIYRCLVG 120

Qy 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
Db 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180

Qy 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEEVAEVEE 240
Db 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEEVAEVEE 240

Qy 241 EADDDDEDEDGEVEEAEPYBEATRTTSIATTTTTTSTESVEEVVREVCSQAETGPC 300
Db 241 EADDDDEDEDGEVEEAEPYBEATRTTSIATTTTTTSTESVEEVVREVCSQAETGPC 300

Qy 301 RAMISRWFVDTGKCAPFFYGGCGGNRNFDTEECYMAVCGSAMQSLLKTTQEPPLARD 360
Db 301 RAMISRWFVDTGKCAPFFYGGCGGNRNFDTEECYMAVCGSAMQSLLKTTQEPPLARD 360

Qy 361 PVKLPPTAASTPDADVXYLETGPDENEHAHFQKAKERLEAKHRMSQVWREWEAERQA 420
Db 361 PVKLPPTAASTPDADVXYLETGPDENEHAHFQKAKERLEAKHRMSQVWREWEAERQA 420

Qy 421 KNLPKADKAVIOHFQKVESLEQEAANERQOLVETHMARVEAMLNDRRLALENYITAL 480
Db 421 KNLPKADKAVIOHFQKVESLEQEAANERQOLVETHMARVEAMLNDRRLALENYITAL 480

Qy 481 QAVPPRPHVFNMLKYYVRAQKDRQHTLKHFEHVRMVDPKKAAQIRSQVWTHLRVIYER 540
Db 481 QAVPPRPHVFNMLKYYVRAQKDRQHTLKHFEHVRMVDPKKAAQIRSQVWTHLRVIYER 540

Qy 541 MNQSLSLYNNPVAEETQDEVELLQEQNYSDVLANMISEPRIYSGNDALMPSLTET 600
Db 541 MNQSLSLYNNPVAEETQDEVELLQEQNYSDVLANMISEPRIYSGNDALMPSLTET 600

Qy 601 KTTVELLPVNGEFSLLDLPQWHSFGADSVPAANTENEPVDPADARGLTTRPGSGLTN 660
Db 601 KTTVELLPVNGEFSLLDLPQWHSFGADSVPAANTENEPVDPADARGLTTRPGSGLTN 660

Qy 661 IKTEISEVKMDAFRHDGSGYEVHVKLVFFAEDVGSNKGAIIGLMVGGVVIATVITL 720

Db 661 IKTEISEVKMDAFRHDGSGYEVHVKLVFFAEDVGSNKGAIIGLMVGGVVIATVITL 720

Qy 721 VMLKKKQYTSIHGVEVDAAVTPERHLSKMQONGYENPTYKFFEQMON 770
Db 721 VMLKKKQYTSIHGVEVDAAVTPERHLSKMQONGYENPTYKFFEQMON 770

RESULT 10
US-09-548-367D-59
; Sequence 59, Application US/09548367D
; Patent No. 6440698
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 29915/6280H
; CURRENT APPLICATION NUMBER: US/09/548,367D
; CURRENT FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 59
; LENGTH: 772
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-548-367D-59

Query Match 100.0%; Score 4058; DB 4; Length 772;
Best Local Similarity 100.0%; Pred. No. 1.1e-280;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLPGLALLLAANTARALEVPTDGNAGLLAEPOIAMFCGRNLNMHNVQNGKWDSPSGTK 60
Db 1 MLPGLALLLAANTARALEVPTDGNAGLLAEPOIAMFCGRNLNMHNVQNGKWDSPSGTK 60

Qy 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTIQNCKRGRKCKOCTHPHFVPIYRCLVG 120
Db 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTIQNCKRGRKCKOCTHPHFVPIYRCLVG 120

Qy 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
Db 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180

Qy 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEEVAEVEE 240
Db 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEEVAEVEE 240

Qy 241 EADDDDEDEDGEVEEAEPYBEATRTTSIATTTTTTSTESVEEVVREVCSQAETGPC 300
Db 241 EADDDDEDEDGEVEEAEPYBEATRTTSIATTTTTTSTESVEEVVREVCSQAETGPC 300

Qy 301 RAMISRWFVDTGKCAPFFYGGCGGNRNFDTEECYMAVCGSAMQSLLKTTQEPPLARD 360
Db 301 RAMISRWFVDTGKCAPFFYGGCGGNRNFDTEECYMAVCGSAMQSLLKTTQEPPLARD 360

Qy 361 PVKLPPTAASTPDADVXYLETGPDENEHAHFQKAKERLEAKHRMSQVWREWEAERQA 420
Db 361 PVKLPPTAASTPDADVXYLETGPDENEHAHFQKAKERLEAKHRMSQVWREWEAERQA 420

Qy 421 KNLPKADKAVIOHFQKVESLEQEAANERQOLVETHMARVEAMLNDRRLALENYITAL 480
Db 421 KNLPKADKAVIOHFQKVESLEQEAANERQOLVETHMARVEAMLNDRRLALENYITAL 480

Qy 481 QAVPPRPHVFNMLKYYVRAEQRQHTLKHFEHVRMVDPKKAAQIRSQVWTHLRVIYER 540
Db 481 QAVPPRPHVFNMLKYYVRAEQRQHTLKHFEHVRMVDPKKAAQIRSQVWTHLRVIYER 540

Db 481 QAVPPRPRHVNMLKKYVRAEQDKRQHTLKHFHVRMVDPKAAQIRSOVMTHLRVIER 540
QY 541 MNQSLSLYNNPVAABEIQDEYDELLQKEQNYSDOVLANMISEPRISYGNDAIMPSTET 600
Db 541 MNQSLSLYNNPVAABEIQDEYDELLQKEQNYSDOVLANMISEPRISYGNDAIMPSTET 600
QY 601 KTTVELLPVNGEFLDDLOPWHISFGADSVDPANTENEVEPVDARPAADRLTTRPGSGLTN 660
Db 601 KTTVELLPVNGEFLDDLOPWHISFGADSVDPANTENEVEPVDARPAADRLTTRPGSGLTN 660
QY 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGALIGLMVGGVVIATVITL 720
Db 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGALIGLMVGGVVIATVITL 720
QY 721 VMLKKKQYTSIHGGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMON 770
Db 721 VMLKKKQYTSIHGGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMON 770

RESULT 11
US-08-123-702-4
; Sequence 4, Application US/08123702
; Patent No. 5604131
; GENERAL INFORMATION:
; APPLICANT: Wadsworth, Samuel
; APPLICANT: Snyder, Benjamin
; APPLICANT: Reddy, Vermuri, B.
; APPLICANT: Wei, Chamer
; TITLE OF INVENTION: A cDNA Genomic Hybrid Sequence Encoding APP770
; Patent No. 5604131
; TITLE OF INVENTION: Containing a Genomic DNA Insert of the KI and OX-2 Regions
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; STREET: 1201 West Peachtree Street
; CITY: Atlanta
; STATE: GA
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 17-SEPT-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: TS1121
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404)-873-8794
; TELEFAX: (404)-873-8795
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 751 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-123-702-4

Query Match 97.2%; Score 3943.5; DB 1; Length 751;
Best Local Similarity 97.4%; Pred. No. 1.4e-272;
Matches 750; Conservative 1; Mismatches 0; Indels 19; Gaps 1;

QY 1: MLPGLALLLAANTARALEVPTDGNAGLLAEPQTAMFCGRNLNMHNMVQNGKWDSPSGTK 60
Db 1 MLPGLALLLAANTARALEVPTDGNAGLLAEPQTAMFCGRNLNMHNMVQNGKWDSPSGTK 60
QY 61 TCIDTKEGILQYCEVYPELQITNVVEANQPVITQNCWKGRKQCKTHPHFVYPCVLG 120

Db 61 TCIDTKEGILQYCEVYPELQITNVVEANQPVITQNCWKGRKQCKTHPHFVYPCVLG 120
QY 121 EFVSDALLVPDKCKFLQERMDYCEHLHWHHTVAKETCSEKSNLNDYGMLLPCGIDKFR 180
Db 121 EFVSDALLVPDKCKFLQERMDYCEHLHWHHTVAKETCSEKSNLNDYGMLLPCGIDKFR 180
QY 181 GVEFVCCPLAEDSDNVSADAEEDSDVWVGADTDYADGSDKVVEVAEEVAEEVEE 240
Db 181 GVEFVCCPLAEDSDNVSADAEEDSDVWVGADTDYADGSDKVVEVAEEVAEEVEE 240
QY 241 EADDEDEDGDEVEEAEPEEATERITTSIATTTTTTTESEVEEVVREYVCSQAETGPC 300
Db 241 EADDEDEDGDEVEEAEPEEATERITTSIATTTTTTTESEVEEVVREYVCSQAETGPC 300
QY 301 RAMLSRWYFDVTEGKCAPFYGGCGNNRNFDEEYCMVCGSAMQSLLKTTQEP LARD 360
Db 301 RAMLSRWYFDVTEGKCAPFYGGCGNNRNFDEEYCMVCGSAMQSLLKTTQEP LARD 360
QY 361 PVKLPTTAASTPDADVDKYLETPGDENEHAFQAKERLEAKHRRMSQVMREWEAEAOA 420
Db 361 PVKLPTTAASTPDADVDKYLETPGDENEHAFQAKERLEAKHRRMSQVMREWEAEAOA 420
QY 421 KNLPKADKKAVIQHFOEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
Db 421 KNLPKADKKAVIQHFOEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
QY 481 QAVPPRPRHVNMLKKYVRAEQDKRQHTLKHFHVRMVDPKAAQIRSOVMTHLRVIER 540
Db 481 QAVPPRPRHVNMLKKYVRAEQDKRQHTLKHFHVRMVDPKAAQIRSOVMTHLRVIER 540
QY 541 MNQSLSLYNNPVAABEIQDEYDELLQKEQNYSDOVLANMISEPRISYGNDAIMPSTET 600
Db 541 MNQSLSLYNNPVAABEIQDEYDELLQKEQNYSDOVLANMISEPRISYGNDAIMPSTET 600
QY 601 KTTVELLPVNGEFLDDLOPWHISFGADSVDPANTENEVEPVDARPAADRLTTRPGSGLTN 660
Db 601 KTTVELLPVNGEFLDDLOPWHISFGADSVDPANTENEVEPVDARPAADRLTTRPGSGLTN 660
QY 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGALIGLMVGGVVIATVITL 720
Db 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGALIGLMVGGVVIATVITL 720
QY 721 VMLKKKQYTSIHGGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMON 770
Db 721 VMLKKKQYTSIHGGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMON 770

RESULT 12
US-08-104-165-2
; Sequence 2, Application US/08104165
; Patent No. 5877015
; GENERAL INFORMATION:
; APPLICANT: HARDY, John Anthony
; APPLICANT: GOATE, Alison Mary
; APPLICANT: MULLAN, Michael John
; APPLICANT: CHARTIER-HARLIN, Marie-Christine
; APPLICANT: OWEN, Michael John
; TITLE OF INVENTION: Test and Model for Alzheimer's Disease
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Khourie and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/104.165

;
; FILING DATE: 21-JAN-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 9101307.8
; FILING DATE: 21-JAN-1991
; APPLICATION NUMBER: 9118445.7
; FILING DATE: 28-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Liebeschuetz, Joe
; REGISTRATION NUMBER: 37,505
; REFERENCE/DOCKET NUMBER: 16163-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 751 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-104-165-2

Query Match 97.2%; Score 3943.5; DB 2; Length 751;
Best Local Similarity 97.4%; Pred. No. 1.4e-272;
Matches 750; Conservative 1; Mismatches 0; Indels 19; Gaps 1;

```
QY 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPOIAMFCGRNLNMHNVQNGKWDSPSGTK 60
Db 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPOIAMFCGRNLNMHNVQNGKWDSPSGTK 60
QY 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTTQNMCKRGKCKOCTHPEHVIYRCLVG 120
Db 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTTQNMCKRGKCKOCTHPEHVIYRCLVG 120
QY 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHTVAKETSEKSTNLHDYGMLLPCGIDKFR 180
Db 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHTVAKETSEKSTNLHDYGMLLPCGIDKFR 180
QY 181 GVEFVCCPLAESDNVDSADAEEDSDVMWGGADTDYADGSEDKVVEVAEEVEAEVEE 240
Db 181 GVEFVCCPLAESDNVDSADAEEDSDVMWGGADTDYADGSEDKVVEVAEEVEAEVEE 240
QY 241 EADDDDEDDGDEVEEAEPEEATERTTSIATTTTTTSEVEEVVRCVSEQAETGPC 300
Db 241 EADDDDEDDGDEVEEAEPEEATERTTSIATTTTTTSEVEEVVRCVSEQAETGPC 300
QY 301 RAMISRWFVDVTEGKCAPFFYGGCGGNRRNFDTEECYMAVCGSAMSSQLLKTTOEPLARD 360
Db 301 RAMISRWFVDVTEGKCAPFFYGGCGGNRRNFDTEECYMAVCGSA----- 344
QY 361 PVKLPPTAASTPDVADKYLETPGDENEHAHFQKAKERLEAKHRRMSQVMRWEAEARQA 420
Db 345 ----IPTTAASTPDVADKYLETPGDENEHAHFQKAKERLEAKHRRMSQVMRWEAEARQA 401
QY 421 KNLPKADKKAIVQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
Db 402 KNLPKADKKAIVQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 461
QY 481 QAVPPRPRHVFNMKKYVRAEOKDRQHTLKHFEHVMVMDPKKAAQIRSQVMTHLRIYER 540
Db 462 QAVPPRPRHVFNMKKYVRAEOKDRQHTLKHFEHVMVMDPKKAAQIRSQVMTHLRIYER 521
QY 541 MNQSLLYNVPAVAEIODEVELLQKEQNYSDDVLANNKISEPRISYGNDAIMPSTLET 600
Db 522 MNQSLLYNVPAVAEIODEVELLQKEQNYSDDVLANNKISEPRISYGNDAIMPSTLET 581
QY 601 KTTVELLPVNGEFLDLDQPHWISFGADSVPAANTEVEPEVDARPAADRLGTLTRPGSLTN 660
Db 582 KTTVELLPVNGEFLDLDQPHWISFGADSVPAANTEVEPEVDARPAADRLGTLTRPGSLTN 641
QY 661 IKTEELSEVKMDAERHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGWGIATVIVITL 720
|||||
```

```
Db 642 IKTEELSEVKMDAERHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGWGIATVIVITL 701
QY 721 VMLKKKQYTSIIHHGVVEVDAAVTPPEERHLSKMQQNGYENPTYKFFEQMON 770
|||||
Db 702 VMLKKKQYTSIIHHGVVEVDAAVTPPEERHLSKMQQNGYENPTYKFFEQMON 751

RESULT 13
US-08-422-333-2
; Sequence 2, Application US/08422333
; Patent No. 5912410
; GENERAL INFORMATION:
; APPLICANT: CORDELL, Barbara L.
; TITLE OF INVENTION: TRANSGENIC NON-HUMAN MAMMAL DISPLAYING
; TITLE OF INVENTION: THE AMYLOID-FORMING PATHOLOGY OF ALZHEIMER'S DISEASE
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scios, Inc.
; STREET: 2450 Bayshore Parkway
; CITY: Mountain View
; STATE: CA
; COUNTRY: USA
; ZIP: 94043
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/422,333
; FILING DATE: 13-APR-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Shearer, Peter R.
; REGISTRATION NUMBER: 28,117
; REFERENCE/DOCKET NUMBER: 21900-28048.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 966-1550
; TELEFAX: (415) 968-2438
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 751 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-422-333-2
```

Query Match 97.2%; Score 3943.5; DB 2; Length 751;
Best Local Similarity 97.4%; Pred. No. 1.4e-272;
Matches 750; Conservative 1; Mismatches 0; Indels 19; Gaps 1;

```
QY 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPOIAMFCGRNLNMHNVQNGKWDSPSGTK 60
Db 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPOIAMFCGRNLNMHNVQNGKWDSPSGTK 60
QY 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTTQNMCKRGKCKOCTHPEHVIYRCLVG 120
Db 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTTQNMCKRGKCKOCTHPEHVIYRCLVG 120
QY 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHTVAKETSEKSTNLHDYGMLLPCGIDKFR 180
Db 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHTVAKETSEKSTNLHDYGMLLPCGIDKFR 180
QY 181 GVEFVCCPLAESDNVDSADAEEDSDVMWGGADTDYADGSEDKVVEVAEEVEAEVEE 240
Db 181 GVEFVCCPLAESDNVDSADAEEDSDVMWGGADTDYADGSEDKVVEVAEEVEAEVEE 240
QY 241 EADDDDEDDGDEVEEAEPEEATERTTSIATTTTTTSEVEEVVRCVSEQAETGPC 300
Db 241 EADDDDEDDGDEVEEAEPEEATERTTSIATTTTTTSEVEEVVRCVSEQAETGPC 300
QY 301 RAMISRWFVDVTEGKCAPFFYGGCGGNRRNFDTEECYMAVCGSAMSSQLLKTTOEPLARD 360
|||||
```

```
Db 301 RAMISRWYDFVTEGKCAPFFYGGCGNRRNFTTEYCMVCGSA----- 344
Qy 361 PVKLPPTAASTPDADVKYLETGPDENEHAFKAKERLEAKHRMSQVMREWEAEARQA 420
Db 345 ---IPTTAASTPDADVKYLETGPDENEHAFKAKERLEAKHRMSQVMREWEAEARQA 401
Qy 421 KNLPKADKAVIOHFQKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITL 480
Db 402 KNLPKADKAVIOHFQKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITL 461
Qy 481 QAVPPRHRVFNMLKYYVRAEQDKRQHTLKHFEHVRMVDPKKAAQIRSQVWTHLRVIYER 540
Db 462 QAVPPRHRVFNMLKYYVRAEQDKRQHTLKHFEHVRMVDPKKAAQIRSQVWTHLRVIYER 521
Qy 541 MNQSLILYNVPAVAEIQDEVDLQKEQNYSDVLANMISEPRISYGNDAIMPSTLSTET 600
Db 522 MNQSLILYNVPAVAEIQDEVDLQKEQNYSDVLANMISEPRISYGNDAIMPSTLSTET 581
Qy 601 KTTVELLPVNGEFSLLDLPWHSGFADSVPAANTENEVEPVDARPAADRGLTTRPGSGLTN 660
Db 582 KTTVELLPVNGEFSLLDLPWHSGFADSVPAANTENEVEPVDARPAADRGLTTRPGSGLTN 641
Qy 661 IKTEISEVMKDAEFRHDSGYEVHHOKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL 720
Db 642 IKTEISEVMKDAEFRHDSGYEVHHOKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL 701
Qy 721 VMLKKKQYTSIIHGVVEVDAVTPPEERHLSKMQQNGYENPTYKFFEQMQN 770
Db 702 VMLKKKQYTSIIHGVVEVDAVTPPEERHLSKMQQNGYENPTYKFFEQMQN 751

RESULT 14
US-08-422-333-21
; Sequence 21. Application US/08422333
; Patent No. 5912410
; GENERAL INFORMATION:
; APPLICANT: CORDELL, Barbara L.
; TITLE OF INVENTION: TRANSGENIC NON-HUMAN MAMMAL DISPLAYING
; THE AMYLOID-FORMING PATHOLOGY OF ALZHEIMER'S DISEASE
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIOS, Inc.
; STREET: 2450 Bayshore Parkway
; CITY: Mountain View
; STATE: CA
; COUNTRY: USA
; ZIP: 94043
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/422,333
; FILING DATE: 13-APR-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Shearer, Peter R.
; REGISTRATION NUMBER: 28,117
; REFERENCE/DOCKET NUMBER: 21900-28048.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 966-1550
; TELEFAX: (415) 968-2438
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 751 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-422-333-21
```

Query Match 97.28; Score 3943.5; DB 2; Length 751;
Best Local Similarity 97.4%; Pred. No. 1.4e-272;

```
Matches 750; Conservative 1; Mismatches 0; Indels 19; Gaps 1;
Qy 1 MLPGLALLLLAAWTARALEVPTDGNAGLLABEQIAMFCGRLNMHNVONGKWDSPSGTK 60
Db 1 MLPGLALLLLAAWTARALEVPTDGNAGLLABEQIAMFCGRLNMHNVONGKWDSPSGTK 60
Qy 61 TCIDTKEGILQYQCEYPELQITNVVEANQPVITQNWCKRGRKQCKTHPHFVPIYRCLVG 120
Db 61 TCIDTKEGILQYQCEYPELQITNVVEANQPVITQNWCKRGRKQCKTHPHFVPIYRCLVG 120
Qy 121 EFVSDALLVPDKCKFLHQBQMDVVCETHLHHWTVAKETCEKSTNLHDYGMLLPCGIDKFR 180
Db 121 EFVSDALLVPDKCKFLHQBQMDVVCETHLHHWTVAKETCEKSTNLHDYGMLLPCGIDKFR 180
Qy 181 GVEFFVCCPLAESDNVDSADAEEDSDVMWGADTDYADGSDKVVEVAEEVEAEVEE 240
Db 181 GVEFFVCCPLAESDNVDSADAEEDSDVMWGADTDYADGSDKVVEVAEEVEAEVEE 240
Qy 241 EADDEDEDEDEGEVEEAEPEYEEATERTTSTATTTTTTSTESVEEVREVCSEQAETGPC 300
Db 241 EADDEDEDEDEGEVEEAEPEYEEATERTTSTATTTTTTSTESVEEVREVCSEQAETGPC 300
Qy 301 RAMISRWYDFVTEGKCAPFFYGGCGNRRNFTTEYCMVCGSAQSLLKTTQEPPLARD 360
Db 301 RAMISRWYDFVTEGKCAPFFYGGCGNRRNFTTEYCMVCGSA----- 344
Qy 361 PVKLPPTAASTPDADVKYLETGPDENEHAFKAKERLEAKHRMSQVMREWEAEARQA 420
Db 345 ---IPTTAASTPDADVKYLETGPDENEHAFKAKERLEAKHRMSQVMREWEAEARQA 401
Qy 421 KNLPKADKAVIOHFQKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITL 480
Db 402 KNLPKADKAVIOHFQKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITL 461
Qy 481 QAVPPRHRVFNMLKYYVRAEQDKRQHTLKHFEHVRMVDPKKAAQIRSQVWTHLRVIYER 540
Db 462 QAVPPRHRVFNMLKYYVRAEQDKRQHTLKHFEHVRMVDPKKAAQIRSQVWTHLRVIYER 521
Qy 541 MNQSLILYNVPAVAEIQDEVDLQKEQNYSDVLANMISEPRISYGNDAIMPSTLSTET 600
Db 522 MNQSLILYNVPAVAEIQDEVDLQKEQNYSDVLANMISEPRISYGNDAIMPSTLSTET 581
Qy 601 KTTVELLPVNGEFSLLDLPWHSGFADSVPAANTENEVEPVDARPAADRGLTTRPGSGLTN 660
Db 582 KTTVELLPVNGEFSLLDLPWHSGFADSVPAANTENEVEPVDARPAADRGLTTRPGSGLTN 641
Qy 661 IKTEISEVMKDAEFRHDSGYEVHHOKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL 720
Db 642 IKTEISEVMKDAEFRHDSGYEVHHOKLVFFAEDVGSNKGAIIGLMVGGVVIATVIVITL 701
Qy 721 VMLKKKQYTSIIHGVVEVDAVTPPEERHLSKMQQNGYENPTYKFFEQMQN 770
Db 702 VMLKKKQYTSIIHGVVEVDAVTPPEERHLSKMQQNGYENPTYKFFEQMQN 751

RESULT 15
US-08-464-250-2
; Sequence 2. Application US/08464250
; Patent No. 6107542
; GENERAL INFORMATION:
; APPLICANT: HARDY, John Anthony
; APPLICANT: GOATE, Alison Mary
; APPLICANT: MULLAN, Michael John
; APPLICANT: CHARTIER-HARLIN, Marie-Christine
; APPLICANT: OWEN, Michael John
; TITLE OF INVENTION: Test and Model for Alzheimer's Disease
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
```

```

; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/464,250
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/104,165
; FILING DATE: 21-JAN-1992
; APPLICATION NUMBER: 9101307.8
; FILING DATE: 21-JAN-1991
; APPLICATION NUMBER: 9118445.7
; FILING DATE: 28-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Liebeschuetz, Joe
; REGISTRATION NUMBER: 37,505
; REFERENCE/DOCKET NUMBER: 16163-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 751 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-464-250-2

```

```

Query Match          97.2%; Score 3943.5; DB 3; Length 751;
Best Local Similarity 97.4%; Pred. No. 1.4e-272;
Matches 750; Conservative 1; Mismatches 0; Indels 19; Gaps 1;

QY 1 MLPLGALLLAATARALEVPTDGNAGLLAEQIAMFCGRNLNMHNQNGKWDSDPSGTK 60
DB 1 MLPLGALLLAATARALEVPTDGNAGLLAEQIAMFCGRNLNMHNQNGKWDSDPSGTK 60
QY 61 TCIDTKEGILQYQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVPIYRCLVG 120
DB 61 TCIDTKEGILQYQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVPIYRCLVG 120
QY 121 EFVSDALLVPDKCKFLHQRMDVCEHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
DB 121 EFVSDALLVPDKCKFLHQRMDVCEHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
QY 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSDKVVEVAEEVEAEVEEE 240
DB 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSDKVVEVAEEVEAEVEEE 240
QY 241 EADDDDEDDGDEVEEAEEPEEATERTSTATTTTTTSEVEEVVREVCSEQAETGPC 300
DB 241 EADDDDEDDGDEVEEAEEPEEATERTSTATTTTTTSEVEEVVREVCSEQAETGPC 300
QY 301 RAMISRWYFDVTGKCAPFFYGGCGNRNNDTEECVMAVCGSAMQSLLKTTQEP LARD 360
DB 301 RAMISRWYFDVTGKCAPFFYGGCGNRNNDTEECVMAVCGSAMQSLLKTTQEP LARD 360
QY 361 PVKLPPTAASDPDAVDKYLETPGDENEHAHFQAKERLEAKHRERMSQVMREWEAEARQA 420
DB 361 PVKLPPTAASDPDAVDKYLETPGDENEHAHFQAKERLEAKHRERMSQVMREWEAEARQA 420
QY 421 KNLPRADKKAVIQHFOEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
DB 421 KNLPRADKKAVIQHFOEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
QY 481 QAVPPPRHVFNNLKKYVRAEQKDRQHTLKHFEHVRMVDPPKAAQIRSQVMTHLRVIER 540
DB 481 QAVPPPRHVFNNLKKYVRAEQKDRQHTLKHFEHVRMVDPPKAAQIRSQVMTHLRVIER 540
QY 541 MNQSLSLLYNPAVAEEIQDEVDELQEQNYSDDVLANMISEPRISYGNDAIMP SLTET 600

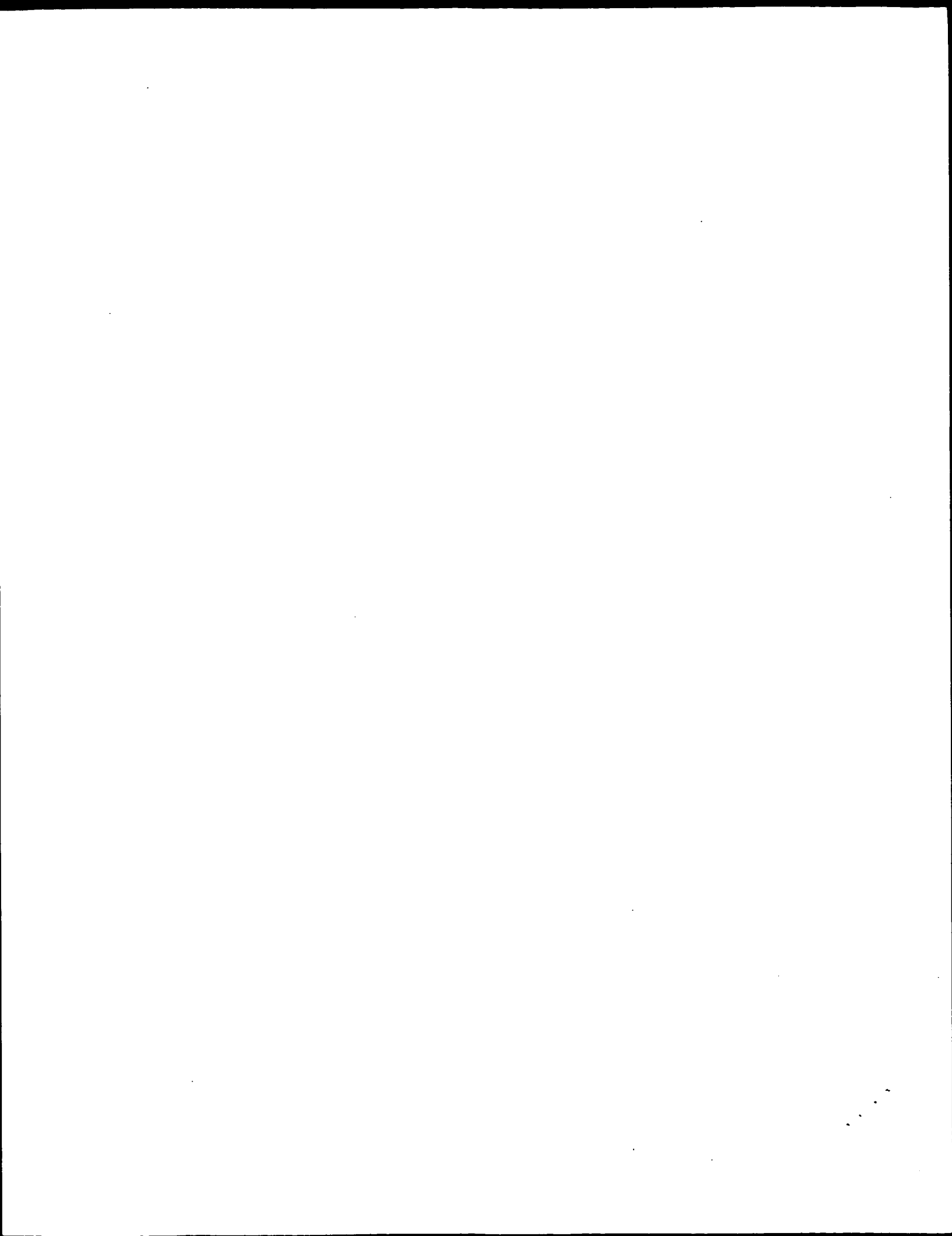
```

Search completed: January 28, 2003, 14:21:20
Job time : 18.9847 secs

```

DB 522 MNQSLSLLYNPAVAEEIQDEVDELQEQNYSDDVLANMISEPRISYGNDAIMP SLTET 581
QY 601 KTTVELLPVNGEFSLDDLPWHSFCADSVPAANTENEVEPVDARPAADRGLTTRPGSGLTN 660
DB 582 KTTVELLPVNGEFSLDDLPWHSFCADSVPAANTENEVEPVDARPAADRGLTTRPGSGLTN 641
QY 661 IKTEEISEVKMDAEPRHDSGYEVHHQKLVFFAEVDVGSNKGAIIGLMVGGVVIATVITL 720
DB 642 IKTEEISEVKMDAEPRHDSGYEVHHQKLVFFAEVDVGSNKGAIIGLMVGGVVIATVITL 701
QY 721 VMLKKKQYTSIHGGVVEVDAAVTPEERHLSKMQQNGYENPTYKPFQEQMN 770
DB 702 VMLKKKQYTSIHGGVVEVDAAVTPEERHLSKMQQNGYENPTYKPFQEQMN 751

```



GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:08:34 ; Search time 6.63425 Seconds
(without alignments)
3666.126 Million cell updates/sec

Title: US-09-904-987-3

Perfect score: 1435

Sequence: 1 MANLGCMVLVLFVATMSDLG.....LFSSPPVILLISFLIFLIVG 253

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1435	100.0	253	1 UJHU	major prion protei
2	1427	99.4	253	2 I61847	major prion protei
3	1427	99.4	253	2 S53635	prion protein - si
4	1427	99.4	253	2 S53614	major prion protei
5	1422	99.1	253	2 S53617	major prion protei
6	1421	99.0	253	2 I37032	major prion protei
7	1409	98.2	253	2 S53616	major prion protei
8	1404	97.8	253	2 S53618	major prion protei
9	1401	97.6	253	2 I84423	major prion protei
10	1401	97.6	253	2 S71055	major prion protei
11	1400	97.6	253	2 S53619	major prion protei
12	1396	97.3	253	2 S53620	major prion protei
13	1396	97.3	253	2 S53623	major prion protei
14	1396	97.3	253	2 S53624	major prion protei
15	1396	97.3	253	2 S53625	major prion protei
16	1386.5	96.6	260	2 S53629	major prion protei
17	1380.5	96.2	252	2 S53634	major prion protei
18	1374.5	95.8	252	2 I61848	major prion protei
19	1370.5	95.5	252	2 S53631	major prion protei
20	1332	92.8	241	2 S71048	major prion protei
21	1331.5	92.8	264	2 A54330	major prion protei
22	1329	92.6	241	2 S71056	major prion protei
23	1329	92.6	245	2 S71045	major prion protei
24	1327	92.5	245	2 S53627	major prion protei
25	1323.5	92.2	254	2 A34759	prion protein - Ch
26	1323.5	92.2	256	2 JU0268	major prion protei
27	1319.5	92.0	254	2 B34759	prion protein - go
28	1316.5	91.7	256	2 A54281	major prion protei
29	1311.5	91.4	256	2 S37149	prion protein - go

RESULT 1

UJHU

major prion protein precursor - human

N:Alternate names: 11K amyloid protein; 27-30K sialoglycoprotein; PrP 27-30; PrP 33-3

C:Species: Homo sapiens (man)

C>Date: 25-Oct-1987 #sequence_revision 12-Apr-1996 #text_change 16-Jun-2000

C:Accession: A24173; A40372; A05017; S14078; I54322; I68597; I58135; I59184; I796333;

R:kretschmar, H.A.; Stowring, L.E.; Westaway, D.; Stubblebine, W.H.; Prusiner, S.B.; DNA 5, 315-324, 1986

A:Title: Molecular cloning of a human prion protein cDNA.

A:Reference number: A24173; MUID:86300093; PMID:3755672

A:Accession: A24173

A:Molecule type: mRNA

A:Residues: 1-253 <KRE>

A:CROSS-references: GB:M13899; NID:g190467; PIDN:AAA60182.1; PID:g190468

R:Puckett, C.; Concannon, P.; Casey, C.; Hood, L.

Am. J. Hum. Genet. 49, 320-329, 1991

A:Title: Genomic structure of the human prion protein gene.

A:Reference number: A40372; MUID:91328137; PMID:1678248

A:Accession: A40372

A:Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-80,89-253 <PUC>

A:CROSS-references: GB:X83416; NID:g747846; PIDN:CAA58442.1; PID:g747847

A:Note: the deletion may be a polymorphism; the alternative deletion of 82-89 could n

R:Liiao, Y.C.J.; Lebo, R.V.; Clawson, G.A.; Smuckler, E.A.

Science 233, 364-367, 1986

A:Reference number: A05017; MUID:86261778; PMID:3014653

A:Accession: A05017

A:Molecule type: mRNA

A:Residues: 8-117,119-253 <LIA>

A:CROSS-references: GB:D00015; NID:g220015; PIDN:BAA00011.1; PID:g220016; GB:M13667;

R:Tagliavini, F.; Prelli, F.; Ghiso, J.; Bugiani, O.; Serban, D.; Prusiner, S.B.; Far

EMBO J. 10, 513-519, 1991

A:Title: Amyloid protein of Gerstmann-Strausler-Scheinker disease (Indiana kindred)

A:Reference number: S14078; MUID:91160504; PMID:1672107

A:Accession: S14078

A:Molecule type: protein

A:Residues: 58-72, 'X', 74-76, 'XX', 79, 'XXX', 83-86:111-128, 'V', 130-150 <TAG>

R:Diedrich, J.F.; Knopman, D.S.; List, J.F.; Olson, K.; Frey, W.H.

Hum. Mol. Genet. 1, 443-444, 1992

A:Title: Deletion in the prion protein gene in a demented patient.

A:Reference number: I54322; MUID:93250789; PMID:1363802

A:Accession: I54322

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 9-83,92-240 <RES>

A:CROSS-references: GB:M81929; NID:g190517; PIDN:AA59442.1; PID:g190518

A:Accession: I68597

A:Status: translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 8-240 <RES3>

prion protein - gr
major prion prp-Sc
major prion protei
major prion protei
major prion protei
major prion prp27-
major prion protei
prion protein - ra
major prion protei
prion-related prot
prion protein - ch
prion protein homo
major prion protei
heterogeneous ribo
glycine-rich cell
glycine-rich prote
keratin, 54K type

ALIGNMENTS


```

Db 1 MANLGCMWLVLVFVATWSDLGLCKKRPKPGWNTGGSRYPGQSPGNGNRYPPQGGGGWGQP 60
QY 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPKTNKHMAGAAAAGA 120
Db 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPKTNKHMAGAAAAGA 120
QY 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENMHRYPNQVYRPMDEYSNQNMFVHDCV 180
Db 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENMHRYPNQVYRPMDEYSNQNMFVHDCV 180
QY 181 NITIKQHTVTTTKGENFTETDVKKMERVVEQMCITQYERESQAYYQRGSSMVLFSPPV 240
Db 181 NITIKQHTVTTTKGENFTETDVKKMERVVEQMCITQYERESQAYYQRGSSMVLFSPPV 240
QY 241 ILLISFLIFLIVG 253
Db 241 ILLISFLIFLIVG 253

RESULT 4
S53614
major prion protein - gorilla
C:Species: Gorilla gorilla (gorilla)
C:Date: 28-Oct-1996 #sequence_revision 07-Feb-1997 #text_change 13-Aug-1999
C:Accession: S53614; S71049
R:Schaetzl, H.M.; da Costa, M.; Taylor, L.; Cohen, F.E.; Prusiner, S.B.
J. Mol. Biol. 245, 362-374, 1995
A:Title: Prion protein gene variation among primates.
A:Reference number: S53614; MUID:95139066; PMID:7837269
A:Accession: S53614
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-253 <SCH>
A:Cross-references: EMBL:U08290
R:Schaetzl, H.M.
submitted to the EMBL Data Library, April 1994
A:Reference number: S71041
A:Accession: S71050
A:Molecule type: DNA
A:Residues: 1-210,'E',212-253 <SCW>
A:Cross-references: EMBL:U08299; NID:g474358; PIDN:AAC50089.1; PID:g474359
C:Superfamily: major prion protein
C:Keywords: amyloid; brain; glycoprotein; lipoprotein; prion; scrapie; transmembrane
protein

Query Match 99.4%; Score 1427; DB 2; Length 253;
Best Local Similarity 99.2%; Pred. No. 4.8e-109;
Matches 251; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MANLGCMWLVLVFVATWSDLGLCKKRPKPGWNTGGSRYPGQSPGNGNRYPPQGGGGWGQP 60
Db 1 MANLGCMWLVLVFVATWSDLGLCKKRPKPGWNTGGSRYPGQSPGNGNRYPPQGGGGWGQP 60
QY 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPKTNKHMAGAAAAGA 120
Db 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPKTNKHMAGAAAAGA 120
QY 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENMHRYPNQVYRPMDEYSNQNMFVHDCV 180
Db 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENMHRYPNQVYRPMDEYSNQNMFVHDCV 180
QY 181 NITIKQHTVTTTKGENFTETDVKKMERVVEQMCITQYERESQAYYQRGSSMVLFSPPV 240
Db 181 NITIKQHTVTTTKGENFTETDVKKMERVVEQMCITQYERESQAYYQRGSSMVLFSPPV 240
QY 241 ILLISFLIFLIVG 253
Db 241 ILLISFLIFLIVG 253

RESULT 5
S53617
major prion protein - common gibbon
C:Species: Hylobates lar (common gibbon, white-handed gibbon)

```

```

C:Date: 14-Feb-1997 #sequence_revision 14-Feb-1997 #text_change 13-Aug-1999
C:Accession: S53617; S71050
R:Schaetzl, H.M.; da Costa, M.; Taylor, L.; Cohen, F.E.; Prusiner, S.B.
J. Mol. Biol. 245, 362-374, 1995
A:Title: Prion protein gene variation among primates.
A:Reference number: S53614; MUID:95139066; PMID:7837269
A:Accession: S53617
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-253 <SCH>
A:Cross-references: EMBL:U08299
R:Schaetzl, H.M.
submitted to the EMBL Data Library, April 1994
A:Reference number: S71041
A:Accession: S71050
A:Molecule type: DNA
A:Residues: 1-210,'E',212-253 <SCW>
A:Cross-references: EMBL:U08299; NID:g474356; PIDN:AAC50088.1; PID:g474357
C:Superfamily: major prion protein
C:Keywords: amyloid; brain; glycoprotein; lipoprotein; prion; scrapie; transmembrane
protein

Query Match 99.1%; Score 1422; DB 2; Length 253;
Best Local Similarity 98.8%; Pred. No. 1.2e-108;
Matches 250; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MANLGCMWLVLVFVATWSDLGLCKKRPKPGWNTGGSRYPGQSPGNGNRYPPQGGGGWGQP 60
Db 1 MANLGCMWLVLVFVATWSDLGLCKKRPKPGWNTGGSRYPGQSPGNGNRYPPQGGGGWGQP 60
QY 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPKTNKHMAGAAAAGA 120
Db 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPKTNKHMAGAAAAGA 120
QY 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENMHRYPNQVYRPMDEYSNQNMFVHDCV 180
Db 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENMHRYPNQVYRPMDEYSNQNMFVHDCV 180
QY 181 NITIKQHTVTTTKGENFTETDVKKMERVVEQMCITQYERESQAYYQRGSSMVLFSPPV 240
Db 181 NITIKQHTVTTTKGENFTETDVKKMERVVEQMCITQYERESQAYYQRGSSMVLFSPPV 240
QY 241 ILLISFLIFLIVG 253
Db 241 ILLISFLIFLIVG 253

RESULT 6
I37032
major prion protein precursor - gorilla
C:Species: Gorilla gorilla (gorilla)
C:Date: 31-May-1996 #sequence_revision 31-May-1996 #text_change 13-Aug-1999
C:Accession: I37032
R:Cervenakova, L.; Brown, P.; Goldfarb, L.G.; Nagle, J.; Petrone, K.; Rubenstein, R.
Proc. Natl. Acad. Sci. U.S.A. 91, 12159-12162, 1994
A:Title: Infectious amyloid precursor gene sequences in primates used for experimenta
A:Reference number: I36907; MUID:95083661; PMID:7991600
A:Accession: I37032
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-253 <RES>
A:Cross-references: EMBL:U15166; NID:g563208; PIDN:AAA68633.1; PID:g563209
C:Superfamily: major prion protein

Query Match 99.0%; Score 1421; DB 2; Length 253;
Best Local Similarity 99.2%; Pred. No. 1.5e-108;
Matches 251; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MANLGCMWLVLVFVATWSDLGLCKKRPKPGWNTGGSRYPGQSPGNGNRYPPQGGGGWGQP 60
Db 1 MANLGCMWLVLVFVATWSDLGLCKKRPKPGWNTGGSRYPGQSPGNGNRYPPQGGGGWGQP 60
QY 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPKTNKHMAGAAAAGA 120

```

Db 61 HGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPTNKKHMAGAAAAGA 120
 QY 121 VVGGLGYMLGSAMSRPIIHFGSDYEDRYRENHMYRPNQVYRPMDEYSNQNNFVHDCV 180
 |||||||
 Db 121 VVGGLGYMLGSAMSRPIIHFGSDYEDRYRENHMYRPNQVYRPMDEYSNQNNFVHDCV 180
 |||||||
 QY 181 NITIKQHTVTTTGTGNTFTDVKMERVVEQMCITQYERESQAYYQRGSSMWLFSSPPV 240
 |||||||
 Db 181 NITIKQHTVTTTGTGNTFTDVKMERVVEQMCITQYERESQAYYQRGSSMWLFSSPPV 240
 |||||||
 QY 241 ILLISFLIFLIVG 253
 |||||||
 Db 241 ILLISFLIFLIVG 253
 |||||||

RESULT 7

S53616

major prion protein - orangutan
 C:Species: Pongo pygmaeus (Orangutan)
 C:Date: 28-Oct-1996 #sequence_revision 07-Feb-1997 #text_change 13-Aug-1999
 C:Accession: S53616; S71059
 R:Schatzl, H.M.; da Costa, M.; Taylor, L.; Cohen, F.E.; Prusiner, S.B.
 J. Mol. Biol. 245, 362-374, 1995
 A:Title: Prion protein gene variation among primates.
 A:Reference number: S53616; MUID:95139066; PMID:7837269
 A:Accession: S53616
 A:Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-253 <SCH>
 A:Cross-references: EMBL:U08305
 R:Schatzl, H.M.
 submitted to the EMBL Data Library, April 1994
 A:Reference number: S71041
 A:Accession: S71059
 A:Molecule type: DNA
 A:Residues: 1-210, 'E', 212-253 <SCW>
 A:Cross-references: EMBL:U08305; NID:g474368; PIDN:AAC50093.1; PID:g474369
 C:Superfamily: major prion protein
 C:Keywords: amyloid; brain; glycoprotein; lipoprotein; prion; scrapie; transmembrane

Query Match 98.2%; Score 1409; DB 2; Length 253;
 Best Local Similarity 97.6%; Pred. No. 1.4e-107;
 Matches 247; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 MANLGCMWLVLVATWSDLGCKKRPKPGWNTGSGRYPGQSPGNGRYPPQGGGWGQP 60
 |||||||
 Db 1 MANLGCMWLVLVATWSNLGCKKRPKPGWNTGSGRYPGQSPGNGRYPPQGGGWGQP 60
 |||||||
 QY 61 HGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPTNKKHMAGAAAAGA 120
 |||||||
 Db 61 HGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPTNKKHMAGAAAAGA 120
 |||||||
 QY 121 VVGGLGYMLGSAMSRPIIHFGSDYEDRYRENHMYRPNQVYRPMDEYSNQNNFVHDCV 180
 |||||||
 Db 121 VVGGLGYMLGSAMSRPIIHFGSDYEDRYRENHMYRPNQVYRPMDEYSNQNNFVHDCV 180
 |||||||

RESULT 8

S53618

major prion protein - Colobus guereza
 C:Species: Colobus guereza
 C:Date: 28-Oct-1996 #sequence_revision 07-Feb-1997 #text_change 13-Aug-1999
 C:Accession: S53618; S71046
 R:Schatzl, H.M.; da Costa, M.; Taylor, L.; Cohen, F.E.; Prusiner, S.B.
 J. Mol. Biol. 245, 362-374, 1995

A:Title: Prion protein gene variation among primates.
 A:Reference number: S53614; MUID:95139066; PMID:7837269
 A:Accession: S53618
 A:Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-253 <SCH>
 A:Cross-references: EMBL:U08297
 R:Schatzl, H.M.
 submitted to the EMBL Data Library, April 1994
 A:Reference number: S71041
 A:Accession: S71046
 A:Molecule type: DNA
 A:Residues: 1-210, 'E', 212-253 <SCW>
 A:Cross-references: EMBL:U08297; NID:g474352; PIDN:AAC50086.1; PID:g474353
 C:Superfamily: major prion protein
 C:Keywords: amyloid; brain; glycoprotein; lipoprotein; prion; scrapie; transmembrane

Query Match 97.8%; Score 1404; DB 2; Length 253;
 Best Local Similarity 96.8%; Pred. No. 3.6e-107;
 Matches 245; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY 1 MANLGCMWLVLVATWSDLGCKKRPKPGWNTGSGRYPGQSPGNGRYPPQGGGWGQP 60
 |||||||
 Db 1 MANLGCMWLVLVATWSDLGCKKRPKPGWNTGSGRYPGQSPGNGRYPPQGGGWGQP 60
 |||||||
 QY 61 HGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPTNKKHMAGAAAAGA 120
 |||||||
 Db 61 HGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPTNKKHMAGAAAAGA 120
 |||||||
 QY 121 VVGGLGYMLGSAMSRPIIHFGSDYEDRYRENHMYRPNQVYRPMDEYSNQNNFVHDCV 180
 |||||||
 Db 121 VVGGLGYMLGSAMSRPIIHFGSDYEDRYRENHMYRPNQVYRPMDEYSNQNNFVHDCV 180
 |||||||
 QY 181 NITIKQHTVTTTGTGNTFTDVKMERVVEQMCITQYERESQAYYQRGSSMWLFSSPPV 240
 |||||||
 Db 181 NITIKQHTVTTTGTGNTFTDVKMERVVEQMCITQYERESQAYYQRGSSMWLFSSPPV 240
 |||||||
 QY 241 ILLISFLIFLIVG 253
 |||||||
 Db 241 ILLISFLIFLIVG 253
 |||||||

RESULT 9

I84423

major prion protein precursor - rhesus macaque
 C:Species: Macaca mulatta (rhesus macaque)
 C:Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 13-Aug-1999
 C:Accession: I84423; S53622; S71054
 R:Cervenakova, L.; Brown, P.; Goldfarb, L.G.; Nagle, J.; Pettrone, K.; Rubenstein, R.
 Proc. Natl. Acad. Sci. U.S.A. 91, 12159-12162, 1994
 A:Title: Infectious amyloid precursor gene sequences in primates used for experimental
 A:Reference number: I36907; MUID:95083661; PMID:7991600
 A:Accession: I84423
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-253 <RES>
 A:Cross-references: EMBL:U15163; NID:g595850; PIDN:AAA68635.1; PID:g595851
 R:Schatzl, H.M.; da Costa, M.; Taylor, L.; Cohen, F.E.; Prusiner, S.B.
 J. Mol. Biol. 245, 362-374, 1995
 A:Title: Prion protein gene variation among primates.
 A:Reference number: S53614; MUID:95139066; PMID:7837269
 A:Accession: S53622
 A:Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-210, 'R', 212-253 <SCH>
 A:Cross-references: EMBL:U08307
 R:Schatzl, H.M.
 submitted to the EMBL Data Library, April 1994
 A:Reference number: S71041
 A:Accession: S71054
 A:Molecule type: DNA
 A:Residues: 1-253 <SCH>
 A:Cross-references: EMBL:U08307; NID:g474372; PIDN:AAC50095.1; PID:g474373

C:Superfamily: major prion protein
C:Keywords: amyloid; brain; glycoprotein; lipoprotein; prion; scrapie; transmembrane protein

Query Match 97.6%; Score 1401; DB 2; Length 253;
Best Local Similarity 96.4%; Pred. No. 6.3e-107;
Matches 244; Conservative 9; Mismatches 0; Indels 0; Gaps 0;

QY 1 MANLGCMVLVLFVATWSDGLCKKRPKPGWNTGSGRYPGQSPGNGRYPPOGGGQGP 60
|||||
Db 1 MANLGCMVLVLFVATWSDGLCKKRPKPGWNTGSGRYPGQSPGNGRYPPOGGGQGP 60
|||||
QY 61 HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP 120
|||||
Db 61 HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP 120
|||||
QY 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNFFVHDCV 180
|||||
Db 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNFFVHDCV 180
|||||
QY 181 NITIKQHTVTTTGTGENTETDVKMERVVEQMCITQYERESQAYYQRGSSMVLFSPPV 240
|||||
Db 181 NITIKQHTVTTTGTGENTETDVKMERVVEQMCITQYERESQAYYQRGSSMVLFSPPV 240
|||||
QY 241 ILLISFLIFLIVG 253
|||||
Db 241 ILLISFLIFLIVG 253
|||||

RESULT 10

S71055

major prion protein - pig-tailed macaque
C:Species: Macaca nemestrina (pig-tailed macaque)
C:Date: 14-Feb-1997 #sequence_revision 14-Feb-1997 #text_change 13-Aug-1999
C:Accession: S71055; S53626

R:Schatzl, H.M.

submitted to the EMBL Data Library, April 1994

A:Reference number: S71041

A:Accession: S71055

A:Molecule type: DNA

A:Residues: 1-253 <SCH>

A:Cross-references: EMBL:U08306; NID:9474370; PIDN:AAC50094.1; PID:9474371

R:Schatzl, H.M.; da Costa, M.; Taylor, L.; Cohen, F.E.; Prusiner, S.B.

J. Mol. Biol. 245, 362-374, 1995

A:Title: Prion protein gene variation among primates.

A:Reference number: S53626; MUID:95139066; PMID:7837269

A:Accession: S53626

A>Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 8-210, R', 212-247 <SCH>

A:Cross-references: EMBL:U08306

C:Superfamily: major prion protein

C:Keywords: amyloid; brain; glycoprotein; lipoprotein; prion; scrapie; transmembrane protein

Query Match 97.6%; Score 1401; DB 2; Length 253;
Best Local Similarity 96.4%; Pred. No. 6.3e-107;
Matches 244; Conservative 9; Mismatches 0; Indels 0; Gaps 0;

QY 1 MANLGCMVLVLFVATWSDGLCKKRPKPGWNTGSGRYPGQSPGNGRYPPOGGGQGP 60
|||||
Db 1 MANLGCMVLVLFVATWSDGLCKKRPKPGWNTGSGRYPGQSPGNGRYPPOGGGQGP 60
|||||
QY 61 HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP 120
|||||
Db 61 HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP 120
|||||
QY 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNFFVHDCV 180
|||||
Db 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNFFVHDCV 180
|||||
QY 181 NITIKQHTVTTTGTGENTETDVKMERVVEQMCITQYERESQAYYQRGSSMVLFSPPV 240
|||||
Db 181 NITIKQHTVTTTGTGENTETDVKMERVVEQMCITQYERESQAYYQRGSSMVLFSPPV 240
|||||

QY 241 ILLISFLIFLIVG 253
|||||
Db 241 ILLISFLIFLIVG 253
|||||

RESULT 11

S53619

major prion protein - Presbytis francoisi

C:Species: Presbytis francoisi

C:Date: 28-Oct-1996 #sequence_revision 07-Feb-1997 #text_change 13-Aug-1999

C:Accession: S53619; S71057

R:Schatzl, H.M.; da Costa, M.; Taylor, L.; Cohen, F.E.; Prusiner, S.B.

J. Mol. Biol. 245, 362-374, 1995

A:Title: Prion protein gene variation among primates.

A:Reference number: S53619; MUID:95139066; PMID:7837269

A:Accession: S53619

A>Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-253 <SCH>

A:Cross-references: EMBL:U08302

R:Schatzl, H.M.

submitted to the EMBL Data Library, April 1994

A:Reference number: S71041

A:Accession: S71057

A:Molecule type: DNA

A:Residues: 1-210, 'E', 212-253 <SCH>

A:Cross-references: EMBL:U08302; NID:gl396067; PIDN:AAB03105.1; PID:gl396068

C:Superfamily: major prion protein

C:Keywords: amyloid; brain; glycoprotein; lipoprotein; prion; scrapie; transmembrane

Query Match 97.6%; Score 1400; DB 2; Length 253;

Best Local Similarity 96.4%; Pred. No. 7.6e-107;

Matches 244; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 MANLGCMVLVLFVATWSDGLCKKRPKPGWNTGSGRYPGQSPGNGRYPPOGGGQGP 60
|||||Db 1 MANLGCMVLVLFVATWSDGLCKKRPKPGWNTGSGRYPGQSPGNGRYPPOGGGQGP 60
|||||QY 61 HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP 120
|||||Db 61 HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP 120
|||||QY 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNFFVHDCV 180
|||||Db 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNFFVHDCV 180
|||||QY 181 NITIKQHTVTTTGTGENTETDVKMERVVEQMCITQYERESQAYYQRGSSMVLFSPPV 240
|||||Db 181 NITIKQHTVTTTGTGENTETDVKMERVVEQMCITQYERESQAYYQRGSSMVLFSPPV 240
|||||QY 241 ILLISFLIFLIVG 253
|||||Db 241 ILLISFLIFLIVG 253
|||||

RESULT 12

S53620

major prion protein - hamadryas baboon

C:Species: Papio hamadryas (hamadryas baboon)

C:Date: 28-Oct-1996 #sequence_revision 07-Feb-1997 #text_change 13-Aug-1999

C:Accession: S53620; S71058

R:Schatzl, H.M.; da Costa, M.; Taylor, L.; Cohen, F.E.; Prusiner, S.B.

J. Mol. Biol. 245, 362-374, 1995

A:Title: Prion protein gene variation among primates.

A:Reference number: S53620; MUID:95139066; PMID:7837269

A:Accession: S53620

A>Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-253 <SCH>

A:Cross-references: EMBL:U08294

R:Schatzl, H.M.

submitted to the EMBL Data Library, April 1994

A:Reference number: S71041

QY 181 NITIKQHTVTTTKGENFTEDVKMMERVVQMCITQYERESQAYYQRGSSMWLFSSPPV 240
Db 181 NITIKQHTVTTTKGENFTEDVKMMERVVQMCITQYERESQAYYQRGSSMWLFSSPPV 240

QY 241 ILLISFLIFLIVG 253
Db 241 ILLISFLIFLIVG 253

RESULT 14
S53624
major prion protein - stump-tailed macaque
C:Species: Macaca arctoides (stump-tailed macaque)
C:Date: 14-Feb-1997 #sequence_revision 14-Feb-1997 #text_change 13-Aug-1999
C:Accession: S53624; S71051
R:Schaetzel, H.M.; da Costa, M.; Taylor, L.; Cohen, F.E.; Prusiner, S.B.
J. Mol. Biol. 245, 362-374, 1995
A:Title: Prion protein gene variation among primates.
A:Reference number: S53614; MUID:95139066; PMID:7837269
A:Accession: S53624
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-253 <SCH>
A:Cross-references: EMBL:U08311
R:Schaetzel, H.M.
submitted to the EMBL Data Library, April 1994
A:Reference number: S71041
A:Accession: S71051
A:Molecule type: DNA
A:Residues: 1-210, 'E', 212-253 <SCW>
A:Cross-references: EMBL:U08311; NID:g475583; PIDN:AAC50099.1; PID:g475584
C:Superfamily: major prion protein
C:Keywords: amyloid; brain; glycoprotein; lipoprotein; prion; scrapie; transmembrane

Query Match 97.38; Score 1396; DB 2; Length 253;
Best Local Similarity 96.08; Pred. No. 1.6e-106;
Matches 243; Conservative 9; Mismatches 1; Indels 0; Gaps 0;

QY 1 MANLGCMWLVFVATWSDGLCKRPGCWNTGGSRYPGGSGPGNRYPPQGSGGWGP 60
Db 1 MANLGCMWLVFVATWSDGLCKRPGCWNTGGSRYPGGSGPGNRYPPQGSGGWGP 60

QY 61 HGGGWGPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPKTKNKHMAAGAAAGA 120
Db 61 HGGGWGPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPKTKNKHMAAGAAAGA 120

QY 121 VVGLGGVYMLGSAMSRPTIHFSGSYEDRYRYRENMHRYPNQVYRPMDEVSNQNNFVHDCV 180
Db 121 VVGLGGVYMLGSAMSRPTIHFSGSYEDRYRYRENMHRYPNQVYRPMDEVSNQNNFVHDCV 180

QY 181 NITIKQHTVTTTKGENFTEDVKMMERVVQMCITQYERESQAYYQRGSSMWLFSSPPV 240
Db 181 NITIKQHTVTTTKGENFTEDVKMMERVVQMCITQYERESQAYYQRGSSMWLFSSPPV 240

QY 241 ILLISFLIFLIVG 253
Db 241 ILLISFLIFLIVG 253

RESULT 15
S53625
major prion protein - Japanese macaque
C:Species: Macaca fuscata (Japanese macaque)
C:Date: 14-Feb-1997 #sequence_revision 14-Feb-1997 #text_change 13-Aug-1999
C:Accession: S53625; S71053
R:Schaetzel, H.M.; da Costa, M.; Taylor, L.; Cohen, F.E.; Prusiner, S.B.
J. Mol. Biol. 245, 362-374, 1995
A:Title: Prion protein gene variation among primates.
A:Reference number: S53614; MUID:95139066; PMID:7837269
A:Accession: S53625
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-253 <SCH>

Search completed: January 28, 2003, 14:20:00
Job time : 7.63425 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:03:18 ; Search time 3.86998 Seconds
(without alignments)
2711.515 Million cell updates/sec

Title: US-09-904-987-3

Perfect score: 1435

Sequence: 1 MANLGCWMLVLPVATWSDLG.....LFSSPPVILLISFLIFLIVG 253

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1435	100.0	253	1 PRIO_HUMAN	P04156 homo sapien
2	1432	99.8	253	1 PRIO_GORGO	P40252 gorilla gor
3	1427	99.4	253	1 PRIO_PANTY	P40253 pan troglod
4	1414	98.5	253	1 PRIO_PONPY	P40256 pongo pygma
5	1409	98.2	253	1 PRIO_COLGU	P40251 colobus gue
6	1405	97.9	253	1 PRIO_PREFR	P40257 presbytis f
7	1401	97.6	253	1 PRIO_MACFA	P40254 macaca fasc
8	1391.5	97.0	260	1 PRIO_SALSC	P40258 salmirus sci
9	1385.5	96.6	252	1 PRIO_CALJA	P40247 callithrix
10	1375.5	95.9	252	1 PRIO_CEBAP	P40249 cebus apell
11	1371.5	95.6	252	1 PRIO_ATEPA	P51446 ateleus pani
12	1356	94.5	246	1 PRIO_CERMO	Q95172 cercopithec
13	1353	94.3	246	1 PRIO_CERPA	Q95174 cercopithec
14	1352	94.2	246	1 PRIO_CERTO	Q95176 cercopithec
15	1331.5	92.8	264	1 PRIO_CALMO	P40248 callicebus
16	1331.5	92.8	264	1 PRIO_BOVIN	P10279 bos taurus
17	1329.5	92.6	254	1 PRIO_SIGHI	Q95023 sigmodon hi
18	1329	92.6	241	1 PRIO_MANSP	P40255 mandrillus
19	1329	92.6	245	1 PRIO_CERAE	P40250 cercopithec
20	1323.5	92.2	254	1 PRIO_CRIGR	Q60506 cricetus
21	1323.5	92.2	256	1 PRIO_BOVIN	Q01880 bos taurus
22	1319.5	92.0	254	1 PRIO_CRIMI	Q60468 cricetus
23	1314.5	91.6	256	1 PRIO_SHEEP	P23907 ovis aries
24	1312.5	91.5	256	1 PRIO_CEREL	P79142 cervus elap
25	1312.5	91.5	256	1 PRIO_ODOHE	P47852 odocolleus
26	1311.5	91.4	256	1 PRIO_CAPHI	P52113 capra hircu
27	1311.5	91.4	264	1 PRIO_PANPY	P40242 tragelaphus
28	1308.5	91.2	254	1 PRIO_MESAU	P40273 mesocricetu
29	1305.5	91.0	239	1 PRIO_AOTTR	P40245 aotus trivi
30	1305	90.9	257	1 PRIO_PIG	P49927 sus scrofa
31	1301.5	90.7	254	1 PRIO_RAT	P13852 rattus norv
32	1299	90.5	257	1 PRIO_MUSVI	P40244 mustela vis
33	1298	90.5	255	1 PRIO_CAMDR	P79141 camelus dro

34	1296.5	90.3	254	1 PRIO_MOUSE	P04925 mus musculu
35	1290.5	89.9	256	1 PRP2_TRAST	P40243 tragelaphus
36	1289	89.8	257	1 PRIO_MUSPF	P52114 mustela put
37	1287	89.7	238	1 PRIO_CERAT	Q95145 cercocebus
38	1280.5	89.2	252	1 PRIO_RABIT	Q95211 oryctolagus
39	1280	89.2	238	1 PRIO_THEGE	Q95270 theropithec
40	1276.5	89.0	256	1 PRIO_FEICA	O18754 felis silve
41	1271	88.6	255	1 PRIO_CANFA	O46501 canis famil
42	1266.5	88.3	232	1 PRIO_ATEGE	P40246 ateleus geof
43	1030	71.8	259	1 PRIO_TRIVU	P51780 trichosurus
44	430.5	30.0	273	1 PRIO_CHICK	P27177 gallus gall
45	173	12.1	465	1 GRP2_PHAVU	P10496 phaseolus v

ALIGNMENTS

RESULT 1

ID	PRIO_HUMAN	STANDARD;	PRT;	253 AA.
AC	P04156;			
DT	01-NOV-1986 (Rel. 03, Created)			
DT	01-NOV-1986 (Rel. 03, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Major prion protein precursor (PrP) (PrP27-30) (PrP33-35C) (ASCR)			
DE	(CD230 antigen).			
GN	PRNP.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=86300093; PubMed=3755672;			
RA	Kretschmar H.A., Stowring L.E., Westaway D., Stubblebine W.H.,			
RA	Prusiner S.B., Dearmond S.J.;			
RT	"Molecular cloning of a human prion protein cDNA.";			
RL	DNA 5:315-324(1986).			
[2]				
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=21638749; PubMed=11780052;			
RA	Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,			
RA	Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguely C.L.,			
RA	Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,			
RA	Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,			
RA	Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,			
RA	Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,			
RA	Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,			
RA	Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,			
RA	Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,			
RA	Graham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,			
RA	Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,			
RA	Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,			
RA	Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,			
RA	Lehvaslahti M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,			
RA	Marsh V.L., Martin S.L., McConachie L.J., McLay K., McMurray A.,			
RA	Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,			
RA	Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,			
RA	Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,			
RA	Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,			
RA	Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,			
RA	Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,			
RA	Tracey A., Romans A.C., Vaudin M., Wall M., Wallis J.M.,			
RA	Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,			
RA	Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,			
Rogers J.;				
"The DNA sequence and comparative analysis of human chromosome 20.";				
Nature 414:865-871(2001).				
[3]				
RP	SEQUENCE OF 8-253 FROM N.A.			
RX	MEDLINE=86261778; PubMed=3014653;			
RA	Liao Y.-C.J., Lebo R.V., Clawson G.A., Smuckler E.A.;			
RT	"Human prion protein cDNA: molecular cloning, chromosomal mapping,			

RT and biological implications.";
 RN Science 233:364-367(1986).
 RA [4]
 RP SEQUENCE OF 58-85 AND 111-150 (VARIANT AMYLOID GSS).
 RX MEDLINE=91160504; PubMed=1672107;
 RA Tagliavini F., Prelli F., Ghiso J., Bugiani O., Serban D.,
 RA Prusiner S.B., Farlow M.R., Ghetti B., Frangione B.;
 RT "Amyloid protein of Gerstmann-Straussler-Scheinker disease (Indiana
 RT kindred) is an 11 kd fragment of prion protein with an N-terminal
 RT glycine at codon 58.";
 RL EMBO J. 10:513-519(1991).
 RN [5]
 RP STRUCTURE BY NMR OF 23-230.
 RX MEDLINE=20087216; PubMed=10618385;
 RA Zahn R., Liu A., Luhrs T., Riek R., von Schroetter C., Wuehrich K.;
 RA Lopez Garcia F., Billeter M., Calzolari L., Wider G., Wuehrich K.;
 RT "NMR solution structure of the human prion protein.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:145-150(2000).
 RN [6]
 RP STRUCTURE BY NMR OF 118-221.
 RX MEDLINE=20359708; PubMed=10900000;
 RA Calzolari L., Lysek D.A., Guntert P., von Schroetter C., Riek R.,
 RA Zahn R., Wuehrich K.;
 RT "NMR structures of three single-residue variants of the human prion
 RT protein.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:8340-8345(2000).
 RN [7]
 RP REVIEW ON VARIANTS.
 RX MEDLINE=93372867; PubMed=8364585;
 RA Palmer M.S., Collinge J.;
 RT "Mutations and polymorphisms in the prion protein gene.";
 RL Hum. Mutat. 2:168-173(1993).
 RN [8]
 RP REVIEW ON VARIANTS.
 RX MEDLINE=94029646; PubMed=8105771;
 RA Prusiner S.B.;
 RT "Genetic and infectious prion diseases.";
 RL Arch. Neurol. 50:1129-1153(1993).
 RN [9]
 RP VARIANT GSS LEU-102.
 RX MEDLINE=89159432; PubMed=2564168;
 RA Hsiao K., Baker H.F., Crow T.J., Poulter M., Owen F.,
 RA Trillinger J.D., Westaway D., Ott J., Prusiner S.B.;
 RT "Linkage of a prion protein missense variant to Gerstmann-Straussler
 RT syndrome.";
 RL Nature 338:342-345(1989).
 RN [10]
 RP VARIANTS LEU-102; VAL-117 AND VAL-129.
 RX MEDLINE=89392018; PubMed=2783132;
 RA Doh-Ura K., Tateishi J., Sasaki H., Kitamoto T., Sakaki Y.;
 RT "Pro-->Leu change at position 102 of prion protein is the most
 RT common but not the sole mutation related to Gerstmann-Straussler
 RT syndrome.";
 RL Biochem. Biophys. Res. Commun. 163:974-979(1989).
 RN [11]
 RP VARIANT FFI ASN-178.
 RX MEDLINE=92195483; PubMed=1347910;
 RA Medori R., Montagna P., Tritschler H.J., Leblanc A., Cortelli P.,
 RA Tinuper P., Lugaresi E., Gambetti P.;
 RT "Fatal familial insomnia: a second kindred with mutation of prion
 RT protein gene at codon 178.";
 RL Neurology 42:669-670(1992).
 RN [12]
 RP VARIANT CJD ASN-178.
 RX MEDLINE=91124933; PubMed=1671440;
 RA Goldfarb L.G., Haltia M., Brown P., Nieto A., Kovanen J.,
 RA McComble W.R., Trapp S., Gajdusek D.C.;
 RT "New mutation in scrapie amyloid precursor gene (at codon 178) in
 RT Finnish Creutzfeldt-Jakob kindred.";
 RL Lancet 337:425-425(1991).
 RN [13]
 RP VARIANT CJD LYS-200.
 RX MEDLINE=90855709; PubMed=1975028;
 RA Goldfarb L., Mitrova E., Brown P., Toh B.K., Gajdusek D.C.;
 RT "Mutation in codon 200 of scrapie amyloid protein gene in two clusters
 RT of Creutzfeldt-Jakob disease in Slovakia.";
 RL Lancet 336:514-515(1990).
 RN [14]
 RP VARIANT GSS ARG-217.
 RX MEDLINE=93250977; PubMed=1363810;
 RA Hsiao K., Dlouhy S.R., Farlow M.R., Cass C., da Costa M.,
 RA Conneally P.M., Hodes M.E., Ghetti B., Prusiner S.B.;
 RT "Mutant prion proteins in Gerstmann-Straussler-Scheinker disease with
 RT neurofibrillary tangles.";
 RL Nat. Genet. 1:68-71(1992).
 RN [15]
 RP VARIANTS CJD ILE-180 AND ARG-232.
 RX MEDLINE=93213314; PubMed=8461023;
 RA Kitamoto T., Ohta M., Doh-Ura K., Hitoshi S., Terao Y., Tateishi J.;
 RT "Novel missense variants of prion protein in Creutzfeldt-Jakob
 RT disease or Gerstmann-Straussler syndrome.";
 RL Biochem. Biophys. Res. Commun. 191:709-714(1993).
 RN [16]
 RP VARIANT CJD ILE-210.
 RX MEDLINE=94071412; PubMed=7902693;
 RA Pochiarri M., Salvatore M., Cutruzzola F., Genuardi M.,
 RA Alcatelli C.T., Masullo C., Macchi G., Alema G., Gaigani S., Xi Y.G.,
 RA Petraroli R., Silvestrini M.C., Brunori M.;
 RT "A new point mutation of the prion protein gene in Creutzfeldt-Jakob
 RT disease.";
 RL Ann. Neurol. 34:802-807(1993).
 RN [17]
 RP VARIANT GSS LEU-105.
 RX MEDLINE=94077414; PubMed=7902972;
 RA Yamada M., Itoh Y., Fujigasaki H., Naruse S., Kaneko K., Kitamoto T.,
 RA Tateishi J., Ohtomo E., Hayakawa M., Tanaka J., Matsushita M.,
 RA Miyatake T.;
 RT "A missense mutation at codon 105 with codon 129 polymorphism of the
 RT prion protein gene in a new variant of Gerstmann-Straussler-Scheinker
 RT disease.";
 RL Neurology 43:2723-2724(1993).
 RN [18]
 RP VARIANT GSS LEU-105.
 RX MEDLINE=95213742; PubMed=7699395;
 RA Itoh Y., Yamada M., Hayakawa M., Shozawa T., Tanaka J., Matsushita M.,
 RA Kitamoto T., Tateishi J., Ohtomo E.;
 RT "A variant of Gerstmann-Straussler-Scheinker disease carrying codon
 RT 105 mutation with codon 129 polymorphism of the prion protein gene: a
 RT clinicopathological study.";
 RL J. Neurol. Sci. 127:77-86(1994).
 RN [19]
 RP VARIANT CJD LYS-200.
 RX MEDLINE=94142912; PubMed=7906019;
 RA Inoue I., Kitamoto T., Doh-Ura K., Shii H., Goto I., Tateishi J.;
 RT "Japanese family with Creutzfeldt-Jakob disease with codon 200 point
 RT mutation of the prion protein gene.";
 RL Neurology 44:299-301(1994).
 RN [20]
 RP VARIANT CJD LYS-200.
 RX MEDLINE=94316708; PubMed=7913755;
 RA Gabizon R., Rosenman H., Meiner Z., Kahana I., Kahana E., Shugart Y.,
 RA Ott J., Prusiner S.B.;
 RT "Mutation in codon 200 and polymorphism in codon 129 of the prion
 RT protein gene in Libyan Jews with Creutzfeldt-Jakob disease.";
 RL Philos. Trans. R. Soc. Lond., B, Biol. Sci. 343:385-390(1994).
 RN [21]
 RP VARIANT GSS LEU-102.
 RX MEDLINE=95303274; PubMed=7783876;
 RA Young K., Jones C.K., Piccardo P., Lazzarini A., Golbe L.I.,
 RA Zimmerman T.R., Dickson D.W., McLachlan D.C., St George-Hyslop P.H.,
 RA Lennox A.;
 RT "Gerstmann-Straussler-Scheinker disease with mutation at codon 102
 RT and methionine at codon 129 of PRNP in previously unreported
 RT patients.";
 RL Neurology 45:1127-1134(1995).
 RN [22]

Query Match 100.0%; Score 1435; DB 1; Length 253;
 Best Local Similarity 100.0%; Pred. No. 6.6e-103;
 Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MANLGCWMLVLFVATWSDLGLCKKRPKPGGNTGGSRYPGQSGPNRYPPQGGGGWGQ 60
 DB 1 MANLGCWMLVLFVATWSDLGLCKKRPKPGGNTGGSRYPGQSGPNRYPPQGGGGWGQ 60

QY 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPTNNKHMAGAAAAGA 120
 DB 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPTNNKHMAGAAAAGA 120

QY 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHMYRPNQVYRPMDEYSNONNFVHDCV 180
 DB 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHMYRPNQVYRPMDEYSNONNFVHDCV 180

QY 181 NITIKQHTVTTTKGENFTETDVKMMERVVEQMCITQYERESQAYYQSGSSMVLFSPPV 240
 DB 181 NITIKQHTVTTTKGENFTETDVKMMERVVEQMCITQYERESQAYYQSGSSMVLFSPPV 240

QY 241 ILLISFLIFLIVG 253
 DB 241 ILLISFLIFLIVG 253

RESULT 2
 ID PRIO_GORGO STANDARD; PRT; 253 AA.
 AC P40252; Q28419;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Major prion protein precursor (PrP) (PrP27-30) (PrP33-35C).
 GN PRNP.
 OS Gorilla gorilla gorilla (Lowland gorilla).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
 OX NCBI_TaxID=9595;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95139066; PubMed=7837269;
 RA Schatzl H.M., Dacosta M., Taylor L., Cohen F.E., Prusiner S.B.;
 RT "Prion protein gene variation among primates.";
 RL J. Mol. Biol. 245:362-374(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Blood;
 RX MEDLINE=95083661; PubMed=7991600;
 RA Cervenakova L., Brown P., Goldfarb L.G., Nagle J., Pettrone K.,
 RA Rubenstein R., Dubnick M., Gibbs C.J., Gajdusek D.C.;
 RT "Infectious amyloid precursor gene sequences in primates used for
 experimental transmission of human spongiform encephalopathy.";
 RL Proc. Natl. Acad. Sci. U.S.A. 91:12159-12162(1994).
 CC -1- FUNCTION: THE FUNCTION OF PRP IS NOT KNOWN. PRP IS ENCODED IN THE
 CC HOST GENOME AND IS EXPRESSED BOTH IN NORMAL AND INFECTED CELLS.
 CC -1- SUBUNIT: PRP HAS A TENDENCY TO AGGREGATE YIELDING POLYMERS CALLED
 CC "RODS".
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
 CC -1- DISEASE: PRP IS FOUND IN HIGH QUANTITY IN THE BRAIN OF HUMANS AND
 CC ANIMALS INFECTED WITH THE DEGENERATIVE NEUROLOGICAL DISEASES KURU,
 CC CREUTZFELDT-JAKOB DISEASE (CJD), GERSTMANN-STRAUSSLER SYNDROME
 CC (GSS), SCRAPIE, BOVINE SPONGIFORM ENCEPHALOPATHY (BSE),
 CC TRANSMISSIBLE MINK ENCEPHALOPATHY (TME), ETC.
 CC -1- SIMILARITY: BELONGS TO THE PRION FAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; U08300; AAC50089.1; -;
 DR EMBL; U15166; AAA68633.1; -;
 DR HSSP; P04156; IQLZ.
 DR InterPro; IPR000817; Prion.
 DR Pfam; PF00377; prion; 1.
 DR PRINTS; PR00341; PRION.
 DR SMART; SM00157; PRP; 1.
 DR PROSITE; PS00291; PRION_1; 1.
 DR PROSITE; PS00706; PRION_2; 1.
 KW Prion; Brain; Glycoprotein; GPI-anchor; Repeat; Signal.
 FT SIGNAL 1 22
 FT CHAIN 23 230
 FT PROPEP 231 253
 FT LIPID 230 230
 FT DISULFID 179 214
 FT CARBOHYD 181 181
 FT CARBOHYD 197 197
 FT DOMAIN 51 91
 FT REPEAT 51 59
 FT REPEAT 60 67
 FT REPEAT 68 75
 FT REPEAT 76 83
 FT REPEAT 84 91
 FT CONFLICT 6 6 C -> Y (IN REF. 2).
 SQ SEQUENCE 253 AA; 27660 MW; E28F4C3FAAEC49E CRC64;

Query Match 99.8%; Score 1432; DB 1; Length 253;
 Best Local Similarity 99.6%; Pred. No. 1.1e-102;
 Matches 252; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MANLGCWMLVLFVATWSDLGLCKKRPKPGGNTGGSRYPGQSGPNRYPPQGGGGWGQ 60
 DB 1 MANLGCWMLVLFVATWSDLGLCKKRPKPGGNTGGSRYPGQSGPNRYPPQGGGGWGQ 60

QY 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPTNNKHMAGAAAAGA 120
 DB 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPTNNKHMAGAAAAGA 120

QY 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHMYRPNQVYRPMDEYSNONNFVHDCV 180
 DB 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHMYRPNQVYRPMDEYSNONNFVHDCV 180

QY 181 NITIKQHTVTTTKGENFTETDVKMMERVVEQMCITQYERESQAYYQSGSSMVLFSPPV 240
 DB 181 NITIKQHTVTTTKGENFTETDVKMMERVVEQMCITQYERESQAYYQSGSSMVLFSPPV 240

QY 241 ILLISFLIFLIVG 253
 DB 241 ILLISFLIFLIVG 253

RESULT 3
 ID PRIO_PANTR STANDARD; PRT; 253 AA.
 AC P40253;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Major prion protein precursor (PrP) (PrP27-30) (PrP33-35C).
 GN PRNP.
 OS Pan troglodytes (Chimpanzee),
 OS Hylobates lar (Common gibbon), and
 OS Hylobates syndactylus (Siamese); (Symphalangus syndactylus).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
 OX NCBI_TaxID=9598, 9580, 9590;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95139066; PubMed=7837269;
 RA Schatzl H.M., Dacosta M., Taylor L., Cohen F.E., Prusiner S.B.;
 RT "Prion protein gene variation among primates.";

DR	InterPro: IPR000817; Prion.
DR	Pfam: PF00377; prion; 1.
DR	PRINTS: PR00341; PRION.
DR	SMART: SM00157; PRP; 1.
DR	PROSITE: PS00291; PRION_1; 1.
DR	PROSITE: PS00706; PRION_2; 1.
KW	Prion; Brain; Glycoprotein; GPI-anchor; Repeat; Signal.
FT	SIGNAL 1 22
FT	CHAIN 23 230
FT	PROPEP 231 253
FT	LIPID 230 230
FT	DISULFID 179 214
FT	CARBOHYD 181 181
FT	CARBOHYD 197 197
FT	DOMAIN 51 91
FT	REPEAT 51 59
FT	REPEAT 60 67
FT	REPEAT 68 75
FT	REPEAT 76 83
FT	REPEAT 84 91
SQ	SEQUENCE 253 AA; 27626 MW; 14B17477881F5316 CRC64;
Query Match 98.2%; Score 1409; DB 1; Length 253;	
Best Local Similarity 97.2%; Pred. No. 6.4e-101;	
Matches 246; Conservative 7; Mismatches 0; Indels 0; Gaps 0;	
QY	1 MANLGCMWLVFVATWSDLGLCKRKPKGWNTGGSRYPGGSPGNRYPPQGGGGWGQP 60
Db	1 MANLGCMWLVFVATWSDLGLCKRKPKGWNTGGSRYPGGSPGNRYPPQGGGGWGQP 60
QY	61 HGGGWGPHGGGWGQPHGGGWGQPHGGGWGQGGTTHSQWNKPSKPKTKNMKHMAGAAAAGA 120
Db	61 HGGGWGPHGGGWGQPHGGGWGQPHGGGWGQGGTTHSQWNKPSKPKTKNMKHMAGAAAAGA 120
QY	121 VVGGLGYMLGSAMSRPIIHFGSDYEDRYRYRENHRYPNQVYYRPMDEYSNQNNFVHDCV 180
Db	121 VVGGLGYMLGSAMSRPLIHFGSDYEDRYRYRENHRYPNQVYYRPMDEYSNQNNFVHDCV 180
QY	181 NITIKQHTVTTTKGENFTEDVKMVRVVEQMCITQYERESQAYYQRGSSMWLFSSPPV 240
Db	181 NITIKQHTVTTTKGENFTEDVKMVRVVEQMCITQYERESQAYYQRGSSMWLFSSPPV 240
QY	241 ILLISFLIFLIVG 253
Db	241 ILLISFLIFLIVG 253
RESULT 6	
ID	PRIO_PREFER STANDARD; PRT; 253 AA.
AC	Q40257;
DT	01-FEB-1995 (Rel. 31, Created)
DT	01-OCT-1996 (Rel. 34, Last sequence update)
DT	01-OCT-1996 (Rel. 34, Last annotation update)
DE	Major prion protein precursor (PrP) (PrP27-30) (PrP33-35C).
GN	PRNP.
OS	Presbytis francoisi.
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae; Colobinae;
OC	Presbytis.
OX	NCBI_TaxID=33549;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=95139066; PubMed=7837269;
RA	Schatzl H.M., Dacosta M., Taylor L., Cohen F.E., Prusiner S.B.;
RT	"Prion protein gene variation among primates..";
RL	J. Mol. Biol. 245:362-374(1995).
CC	-!- FUNCTION: THE FUNCTION OF PRP IS NOT KNOWN. PRP IS ENCODED IN THE
CC	HOST GENOME AND IS EXPRESSED BOTH IN NORMAL AND INFECTED CELLS.
CC	-!- SUBUNIT: PRP HAS A TENDENCY TO AGGREGATE YIELDING POLYMERS CALLED
CC	"RODS".
CC	-!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.

CC -1- DISEASE: PRP IS FOUND IN HIGH QUANTITY IN THE BRAIN OF HUMANS AND
 CC ANIMALS INFECTED WITH THE DEGENERATIVE NEUROLOGICAL DISEASES KURU,
 CC CREUTZFELDT-JAKOB DISEASE (CJD), GERSTMANN-STRAUSSLER SYNDROME
 CC (GSS), SCRAPIE, BOVINE SPONGIFORM ENCEPHALOPATHY (BSE),
 CC TRANSMISSIBLE MINK ENCEPHALOPATHY (TME), ETC.
 CC -1- SIMILARITY: BELONGS TO THE PRION FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U08302; AAB03105.1; -
 DR HSSP; P04925; 1AG2.
 DR InterPro: IPR000817; Prion.
 DR Pfam: PF00377; prion; 1.
 DR PRINTS; PR00341; PRION.
 DR SMART; SM00157; PRP; 1.
 DR PROSITE; PS00291; PRION_1; 1.
 DR PROSITE; PS00706; PRION_2; 1.
 KW Prion; Brain; Glycoprotein; GPI-anchor; Repeat; Signal.
 FT SIGNAL 1 22
 FT CHAIN 23 230 MAJOR PRION PROTEIN.
 FT PROPEP 231 253 REMOVED IN MATURE FORM (BY SIMILARITY).
 FT LIPID 230 230 GPI-ANCHOR (BY SIMILARITY).
 FT DISULFID 179 214 BY SIMILARITY.
 FT CARBOHYD 181 181 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT DOMAIN 51 91 5 X 8 AA TANDEM REPEATS OF P-H-G-G-G-W-G-
 Q.
 FT REPEAT 51 59 1.
 FT REPEAT 60 67 2.
 FT REPEAT 68 75 3.
 FT REPEAT 76 83 4.
 FT REPEAT 84 91 5.
 SQ SEQUENCE 253 AA; 27660 MW; 14B1747AE57F5316 CRC64;
 Query Match 97.9%; Score 1405; DB 1; Length 253;
 Best Local Similarity 96.8%; Pred. No. 1.3e-100;
 Matches 245; Conservative 7; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MANLGCMWLFLVATWSDLGLCKRPRPGGNTGGSRYPGQSPGNGRYPQGGGGWGQP 60
 DB 1 MANLGCMWLFLVATWSDLGLCKRPRPGGNTGGSRYPGQSPGNGRYPQGGGGWGQP 60
 QY 61 HGGGWPQHGCGWQPHGGGQWQPHGGGQGGGTHSQWNKPSKPKTNMKHMAGAAAAGA 120
 DB 61 HGGGWPQHGCGWQPHGGGQWQPHGGGQGGGTHSQWNKPSKPKTNMKHMAGAAAAGA 120
 QY 121 VVGGLGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNNFVHDCV 180
 DB 121 VVGGLGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNNFVHDCV 180
 QY 181 NITIKQHTVTTTNGENFTEDVKMRRVVEQMCITQYERESQAYYQRGSSWVLFSSPPV 240
 DB 181 NITIKQHTVTTTNGENFTEDVKMRRVVEQMCITQYERESQAYYQRGSSWVLFSSPPV 240
 QY 241 ILLISFLIFLIVG 253
 DB 241 ILLISFLIFLIVG 253
 RESULT 7
 ID PRIO_MACFA STANDARD; PRP; 253 AA.
 AC P40254;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Major prion protein precursor (PrP) (PrP27-30) (PrP33-35C).

GN PRNP.
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey),
 OS Macaca arctoides (Stump-tailed macaque),
 OS Macaca fuscata fuscata (Japanese macaque),
 OS Macaca mulatta (Rhesus macaque),
 OS Macaca nemestrina (Pig-tailed macaque), and
 OS Papio hamadryas (Hamadryas baboon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Macaca.
 OX NCBI_TaxID=9541, 9540, 9543, 9544, 9545, 9557;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLIN=95139066; PubMed=7837269;
 RX Schatzl H.M., Dacosta M., Taylor L., Cohen F.E., Prusiner S.B.;
 RT "Prion protein gene variation among primates.";
 RL J. Mol. Biol. 245:362-374(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=M.mulatta; TISSUE=Brain;
 RX MEDLIN=95083661; PubMed=7991600;
 RA Cervenakova L., Brown P., Goldfarb L.G., Nagle J., Pettrone K.,
 RA Rubenstein R., Dubnick M., Gibbs C.J., Gajdusek D.C.;
 RT "Infectious amyloid precursor gene sequences in primates used for
 experimental transmission of human spongiform encephalopathy.";
 RL Proc. Natl. Acad. Sci. U.S.A. 91:12159-12162(1994).
 CC -1- FUNCTION: THE FUNCTION OF PRP IS NOT KNOWN. PRP IS ENCODED IN THE
 CC HOST GENOME AND IS EXPRESSED BOTH IN NORMAL AND INFECTED CELLS.
 CC -1- SUBUNIT: PRP HAS A TENDENCY TO AGGREGATE YIELDING POLYMERS CALLED
 CC "RODS".
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
 CC -1- DISEASE: PRP IS FOUND IN HIGH QUANTITY IN THE BRAIN OF HUMANS AND
 CC ANIMALS INFECTED WITH THE DEGENERATIVE NEUROLOGICAL DISEASES KURU,
 CC CREUTZFELDT-JAKOB DISEASE (CJD), GERSTMANN-STRAUSSLER SYNDROME
 CC (GSS), SCRAPIE, BOVINE SPONGIFORM ENCEPHALOPATHY (BSE),
 CC TRANSMISSIBLE MINK ENCEPHALOPATHY (TME), ETC.
 CC -1- SIMILARITY: BELONGS TO THE PRION FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U08298; AAC50087.1; -
 DR EMBL; U08311; AAC50099.1; -
 DR EMBL; U08301; AAC50090.1; -
 DR EMBL; U08307; AAC50095.1; -
 DR EMBL; U08306; AAC50094.1; -
 DR EMBL; U08294; AAC50083.1; -
 DR EMBL; U15163; AAA68635.1; -
 DR HSSP; P04925; 1AG2.
 DR InterPro: IPR000817; Prion.
 DR Pfam: PF00377; prion; 1.
 DR PRINTS; PR00341; PRION.
 DR SMART; SM00157; PRP; 1.
 DR PROSITE; PS00291; PRION_1; 1.
 DR PROSITE; PS00706; PRION_2; 1.
 KW Prion; Brain; Glycoprotein; GPI-anchor; Repeat; Signal.
 FT SIGNAL 1 22
 FT CHAIN 23 230 MAJOR PRION PROTEIN.
 FT PROPEP 231 253 REMOVED IN MATURE FORM (BY SIMILARITY).
 FT LIPID 230 230 GPI-ANCHOR (BY SIMILARITY).
 FT DISULFID 179 214 BY SIMILARITY.
 FT CARBOHYD 181 181 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT DOMAIN 51 91 5 X 8 AA TANDEM REPEATS OF P-H-G-G-G-W-G-
 Q.
 FT REPEAT 51 59 1.
 FT REPEAT 60 67 2.
 FT REPEAT 68 75 3.


```

CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; U75386; AAB50625.1; -.
CC EMBL; U75387; AAB50626.1; -.
CC HSSP; P04925; IAG2.
CC InterPro; IPR000817; Prion.
CC Pfam; PF00377; prion; 1.
CC SMART; SM00157; prp; 1.
CC PROSITE; PS00291; PRION_1; 1.
CC PROSITE; PS00706; PRION_2; 1.
CC Prion; Brain; Glycoprotein; GPI-anchor; Repeat; Signal.
CC NON_TER 1 1
CC SIGNAL <1 15
CC CHAIN 16 23 MAJOR PRION PROTEIN.
CC PROPEP 224 246 REMOVED IN MATURE FORM (BY SIMILARITY).
CC LIPID 223 223 GPI-ANCHOR (BY SIMILARITY).
CC DISULFID 172 207 BY SIMILARITY.
CC CARBOHYD 174 174 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 190 190 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC DOMAIN 44 84 5 X 8 AA TANDEM REPEATS OF P-H-G-G-G-W-G-
CC REPEAT 44 52 Q.
CC REPEAT 53 60 1.
CC REPEAT 61 68 2.
CC REPEAT 69 76 3.
CC REPEAT 77 84 4.
CC REPEAT 77 84 5.
CC SEQUENCE 246 AA; 26900 MW; 835D147CA2B4FDD3 CRC64;
CC
CC Query Match 94.5%; Score 1356; DB 1; Length 246;
CC Best Local Similarity 96.3%; Pred. No. 6.9e-97;
CC Matches 237; Conservative 9; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 8 MLVLFVATWSDLGLCKRPPKPGGNTGGSRYPGQSPGNGNRYPPQGGGQWGPQHGQ 67
CC DB 1 MLVLFVATWSDLGLCKRPPKPGGNTGGSRYPGQSPGNGNRYPPQGGGQWGPQHGQ 60
CC
CC QY 68 PHGGGQWGPQHGQWGPQHGQWGPQHGQWGPQHGQWGPQHGQWGPQHGQWGPQHGQ 127
CC DB 61 PHGGGQWGPQHGQWGPQHGQWGPQHGQWGPQHGQWGPQHGQWGPQHGQWGPQHGQ 120
CC
CC QY 128 YMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNHFVHDCVNTIKOH 187
CC DB 121 YMLGSAMSRPLIHFGNDYEDRYRENHRYPNQVYRPMDEYSNQNHFVHDCVNTIKOH 180
CC
CC QY 188 TVTTTNGENFTETDVKMERVVEQMCITQYERESQAYYQSGSMVLFSSPPVILLISFL 247
CC DB 181 TVTTTNGENFTETDVKMERVVEQMCITQYERESQAYYQSGSMVLFSSPPVILLISFL 240
CC
CC QY 248 IFLIVG 253
CC DB 241 IFLIVG 246
CC
CC RESULT 13
CC PRIO_CERPA STANDARD; PRT; 246 AA.
CC AC Q95174;
CC DT 01-NOV-1997 (Rel. 35, Created)
CC DT 01-NOV-1997 (Rel. 35, Last sequence update)
CC DT 01-NOV-1997 (Rel. 35, Last annotation update)
CC DE Major prion protein precursor (PrP) (PrP27-30) (PrP33-35C) (Fragment).
CC GN PRNP.
CC OS Cercopithecus patas.
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
CC OC Cercopithecinae; Cercopithecus.
CC OX NCBI_TaxID=27677;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RA der Kuyil A.C., Dekker J.T., Goudsmit J.;
CC RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC CC -!- FUNCTION: THE FUNCTION OF PRP IS NOT KNOWN. PRP IS ENCODED IN THE
CC CC HOST GENOME AND IS EXPRESSED BOTH IN NORMAL AND INFECTED CELLS.

```

```

CC -!- SUBUNIT: PRP HAS A TENDENCY TO AGGREGATE YIELDING POLYMERS CALLED
CC "RODS".
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -!- DISEASE: PRP IS FOUND IN HIGH QUANTITY IN THE BRAIN OF HUMANS AND
CC ANIMALS INFECTED WITH THE DEGENERATIVE NEUROLOGICAL DISEASES KURU,
CC CREUTZFELDT-JAKOB DISEASE (CJD), GERSTMANN-STRAUSSLER SYNDROME
CC (GSS), SCRAPIE, BOVINE SPONGIFORM ENCEPHALOPATHY (BSE),
CC TRANSMISSIBLE MINK ENCEPHALOPATHY (TME), ETC.
CC -!- SIMILARITY: BELONGS TO THE PRION FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; U75388; AAB50627.1; -.
CC HSSP; P04925; IAG2.
CC InterPro; IPR000817; Prion.
CC Pfam; PF00377; prion; 1.
CC SMART; SM00157; prp; 1.
CC PROSITE; PS00291; PRION_1; 1.
CC PROSITE; PS00706; PRION_2; 1.
CC Prion; Brain; Glycoprotein; GPI-anchor; Repeat; Signal.
CC NON_TER 1 1
CC SIGNAL <1 15 BY SIMILARITY.
CC CHAIN 16 23 MAJOR PRION PROTEIN.
CC PROPEP 224 246 REMOVED IN MATURE FORM (BY SIMILARITY).
CC LIPID 223 223 GPI-ANCHOR (BY SIMILARITY).
CC DISULFID 172 207 BY SIMILARITY.
CC CARBOHYD 174 174 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 190 190 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC DOMAIN 44 84 5 X 8 AA TANDEM REPEATS OF P-H-G-G-G-W-G-
CC REPEAT 44 52 Q.
CC REPEAT 53 60 1.
CC REPEAT 61 68 2.
CC REPEAT 69 76 3.
CC REPEAT 77 84 4.
CC REPEAT 77 84 5.
CC SEQUENCE 246 AA; 26886 MW; D35D105BBEC53108 CRC64;
CC
CC Query Match 94.3%; Score 1353; DB 1; Length 246;
CC Best Local Similarity 95.9%; Pred. No. 1.2e-96;
CC Matches 236; Conservative 10; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 8 MLVLFVATWSDLGLCKRPPKPGGNTGGSRYPGQSPGNGNRYPPQGGGQWGPQHGQ 67
CC DB 1 MLVLFVATWSDLGLCKRPPKPGGNTGGSRYPGQSPGNGNRYPPQGGGQWGPQHGQ 60
CC
CC QY 68 PHGGGQWGPQHGQWGPQHGQWGPQHGQWGPQHGQWGPQHGQWGPQHGQWGPQHGQ 127
CC DB 61 PHGGGQWGPQHGQWGPQHGQWGPQHGQWGPQHGQWGPQHGQWGPQHGQWGPQHGQ 120
CC
CC QY 128 YMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNHFVHDCVNTIKOH 187
CC DB 121 YMLGSAMSRPLIHFGNDYEDRYRENHRYPNQVYRPMDEYSNQNHFVHDCVNTIKOH 180
CC
CC QY 188 TVTTTNGENFTETDVKMERVVEQMCITQYERESQAYYQSGSMVLFSSPPVILLISFL 247
CC DB 181 TVTTTNGENFTETDVKMERVVEQMCITQYERESQAYYQSGSMVLFSSPPVILLISFL 240
CC
CC QY 248 IFLIVG 253
CC DB 241 IFLIVG 246
CC
CC RESULT 14
CC PRIO_CERPA STANDARD; PRT; 246 AA.
CC AC Q95176;
CC DT 01-NOV-1997 (Rel. 35, Created)

```

DT 01-Nov-1997 (Rel. 35, Last sequence update)
DT 01-Nov-1997 (Rel. 35, Last annotation update)
DE Major prion protein precursor (PrP) (PrP27-30) (PrP33-35C) (Fragment).
GN PRNP.
OS Cercopithecus torquatus atys (Red-crowned mangabey) (Sooty mangabey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Cercocebus.
ON NCBI_TaxID=9531;
RN [1]
RP SEQUENCE FROM N.A.
RA der Kuyil A.C., Dekker J.T., Goudsmit J.;
RL Submitted (Nov-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: THE FUNCTION OF PRP IS NOT KNOWN. PRP IS ENCODED IN THE
CC HOST GENOME AND IS EXPRESSED BOTH IN NORMAL AND INFECTED CELLS.
CC -!- SUBUNIT: PRP HAS A TENDENCY TO AGGREGATE YIELDING POLYMERS CALLED
CC "RODS".
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -!- DISEASE: PRP IS FOUND IN HIGH QUANTITY IN THE BRAIN OF HUMANS AND
CC ANIMALS INFECTED WITH THE DEGENERATIVE NEUROLOGICAL DISEASES KURU,
CC CREUTZFELDT-JAKOB DISEASE (CJD), GERSTMANN-STRAUSSLER SYNDROME
CC (GSS), SCRAPIE, BOVINE SPONGIFORM ENCEPHALOPATHY (BSE),
CC TRANSMISSIBLE MINK ENCEPHALOPATHY (TME), ETC.
CC -!- SIMILARITY: BELONGS TO THE PRION FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U75385; AAB50628.1; -.
CC HSSP: P04925; IAG2.
CC InterPro: IPR000817; Prion.
CC Pfam: PF00377; prion; 1.
CC SMART: SM00157; PRP; 1.
CC PROSITE: PS00291; PRION_1; 1.
CC PROSITE: PS00706; PRION_2; 1.
CC Prion; Brain; Glycoprotein; GPI-anchor; Repeat; Signal.
CC NON_TER 1 1
CC SIGNAL <1 15
CC CHAIN 16 223
CC PROPEP 224 246
CC LIPID 223 223
CC DIPEPT 223 207
CC DISULFID 172 207
CC CARBOHYD 174 174
CC CARBOHYD 190 190
CC DOMAIN 44 84
CC REPEAT 44 52
CC REPEAT 53 60
CC REPEAT 61 68
CC REPEAT 69 76
CC REPEAT 77 84
CC SEQUENCE 246 AA; 26914 MW; F58679CBCE5ADC7 CRC64;
Query Match 94.2%; Score 1352; DB 1; Length 246;
Best Local Similarity 95.9%; Pred. No. 1.4e-96;
Matches 236; Conservative 10; Mismatches 0; Indels 0; Gaps 0;
QY 8 MLVLFVATWSDGLCKRKPCKGWNTPGQSGPGNRYPPQGGGSGWGPQGGGQ 67
Db 1 MLVLFVATWSDGLCKRKPCKGWNTPGQSGPGNRYPPQGGGSGWGPQGGGQ 60
QY 68 PHGGGWGQPHGGGWGQPHGGGWGQGGGTHSNQWPKTKNKHMAAGAAVVGGLG 127
Db 61 PHGGGWGQPHGGGWGQPHGGGWGQGGGTHSNQWPKTKNKHMAAGAAVVGGLG 120
QY 128 YMLGSAMSRPLIHFGSDYEDRYRENHRYPNQVYRPMDEYSNNQNFVHDCVNTIKQH 187
Db 121 YMLGSAMSRPLIHFGNEYEDRYRENHRYPNQVYRPMDEYSNNQNFVHDCVNTIKQH 180

QY 188 TVTTTTCGENTETDVKMMERVVQMCITQYERESQAYYQRGSSMVLFSPPVILLISFL 247
Db 181 TVTTTTCGENTETDVKMMERVVQMCITQYERESQAYYQRGSSMVLFSPPVILLISFL 240
QY 248 IFLVIG 253
Db 241 IFLVIG 246
RESULT 15
ID PRIO_CALMO STANDARD; PRT; 241 AA.
AC P40248;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Major prion protein precursor (PrP) (PrP27-30) (PrP33-35C) (Fragment).
GN PRNP.
OS Callicebus moloch (Dusky titi).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Callicebinae;
OC Callicebus.
ON NCBI_TaxID=9523;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95139066; PubMed=7837269;
RA Schatzl H.M., Dacosta M., Taylor L., Cohen F.E., Prusiner S.B.;
RL "Prion protein gene variation among primates.";
RL J. Mol. Biol. 245:362-374(1995).
CC -!- FUNCTION: THE FUNCTION OF PRP IS NOT KNOWN. PRP IS ENCODED IN THE
CC HOST GENOME AND IS EXPRESSED BOTH IN NORMAL AND INFECTED CELLS.
CC -!- SUBUNIT: PRP HAS A TENDENCY TO AGGREGATE YIELDING POLYMERS CALLED
CC "RODS".
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -!- DISEASE: PRP IS FOUND IN HIGH QUANTITY IN THE BRAIN OF HUMANS AND
CC ANIMALS INFECTED WITH THE DEGENERATIVE NEUROLOGICAL DISEASES KURU,
CC CREUTZFELDT-JAKOB DISEASE (CJD), GERSTMANN-STRAUSSLER SYNDROME
CC (GSS), SCRAPIE, BOVINE SPONGIFORM ENCEPHALOPATHY (BSE),
CC TRANSMISSIBLE MINK ENCEPHALOPATHY (TME), ETC.
CC -!- SIMILARITY: BELONGS TO THE PRION FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U08312; AAC50100.1; -.
CC HSSP: P04925; IAG2.
CC InterPro: IPR000817; Prion.
CC Pfam: PF00377; prion; 1.
CC SMART: SM00157; PRP; 1.
CC PROSITE: PS00291; PRION_1; 1.
CC PROSITE: PS00706; PRION_2; 1.
CC Prion; Brain; Glycoprotein; GPI-anchor; Repeat; Signal.
CC NON_TER 1 1
CC SIGNAL <1 15
CC CHAIN 16 >241
CC DISULFID 172 207
CC CARBOHYD 174 174
CC CARBOHYD 190 190
CC DOMAIN 44 84
CC REPEAT 44 52
CC REPEAT 53 60
CC REPEAT 61 68
CC REPEAT 69 76
CC REPEAT 77 84
CC NON_TER 241 241
CC SEQUENCE 241 AA; 26373 MW; C6D2013EE7CAEC93 CRC64;

Best Local Similarity 88.8%; Pred. No. 8.3e-122;
Matches 253; Conservative 0; Mismatches 0; Indels 32; Gaps 1;

QY 1 MANLGCMVLVLFVATWSDGLCKRKPXPGGNTGSGRYPGQSPGNGRYPPOG-----53
DB 1 MANLGCMVLVLFVATWSDGLCKRKPXPGGNTGSGRYPGQSPGNGRYPPOGSGWGQP 60
QY 54 -----GGGQPHGGGQPHGGGQPHGGGQPHGGGQPHGGG 88
DB 61 HGGGQPHGGGQPHGGGQPHGGGQPHGGGQPHGGGQPHGGGQPHGGG 120
QY 89 WQGGGTHSQNKKPSKPTNNKHAGAAAAGAVVGGGLGGMGSAMSRPIIHFGSDYEDR 148
DB 121 WQGGGTHSQNKKPSKPTNNKHAGAAAAGAVVGGGLGGMGSAMSRPIIHFGSDYEDR 180
QY 149 YYRENMHRYPNQVYRPMDEYSQNNFVHDCVNIHQHTVTTTKGENFTETDVKMMER 208
DB 181 YYRENMHRYPNQVYRPMDEYSQNNFVHDCVNIHQHTVTTTKGENFTETDVKMMER 240
QY 209 VVEQMCITQYERESQAYYQSGSSMVLFSPPVILLISFLIFLIVG 253
DB 241 VVEQMCITQYERESQAYYQSGSSMVLFSPPVILLISFLIFLIVG 285

RESULT 5
O60489 PRELIMINARY; PRT; 246 AA.
AC O60489;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Prion protein variant (Fragment).
GN PRNP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98044028; PubMed=9384372;
RA Samia H.B., Mari J.J., Vallada H.P., Moura R.P., Simpson A.J.,
Brentani R.R.;
RT "A prion-linked psychiatric disorder.";
RL Nature 390:241-241(1997).
DR EMBL: AF030575; AAC05365.1; -;
DR HSP: P04156; IQLZ.
DR InterPro: IPR000817; Prion.
DR Pfam: PF00377; prion; 1.
DR PRINTS: PR00341; PRION.
DR SMART: SM00157; PRP; 1.
DR PROSITE: PS00291; PRION_1; 1.
DR PROSITE: PS00706; PRION_2; 1.
FT NON_TER 1
SQ SEQUENCE 246 AA; 26826 MW; 1D9B122EA7D1C18C CRC64;

Query Match 96.2%; Score 1381; DB 4; Length 246;
Best Local Similarity 99.2%; Pred. No. 2.6e-119;
Matches 244; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 8 MLVLFVATWSDGLCKRKPXPGGNTGSGRYPGQSPGNGRYPPOGSGWGQP 67
DB 1 MLVLFVATWSDGLCKRKPXPGGNTGSGRYPGQSPGNGRYPPOGSGWGQP 60

QY 68 PHGGGQPHGGGQPHGGGQPHGGGQPHGGGQPHGGGQPHGGGQPHGGG 127
DB 61 PHGGGQPHGGGQPHGGGQPHGGGQPHGGGQPHGGGQPHGGGQPHGGG 120

QY 128 YMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSQNNFVHDCVNIHQHTVTTTKQH 187
DB 121 YVLSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSQNNFVHDCVNIHQHTVTTTKQH 180

QY 188 TVTTTGTGENTETDVKMMERVVEQMCITQYERESQAYYQSGSSMVLFSPPVILLISFL 247
DB 188 TVTTTGTGENTETDVKMMERVVEQMCITQYERESQAYYQSGSSMVLFSPPVILLISFL 247

DB 181 TVTTTGTGENTETDVKMMERVVEQMCITQYERESQAYYQSGSSMVLFSPPVILLISFL 240
QY 248 IFLIVG 253
DB 241 IFLIVG 246

RESULT 6
Q15216 PRELIMINARY; PRT; 245 AA.
AC Q15216; Q15221;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Prion protein.
GN PRP OR PRNP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RX MEDLINE=91328137; PubMed=1678248;
RA Puckett C., Concannon P., Casey C., Hood L.;
RT "Genomic structure of the human prion protein gene.";
RL Am. J. Hum. Genet. 49:320-329(1991).
RN [2]
RP SEQUENCE FROM N.A.
RA Lee I.Y., Westaway D., Smit A.F., Wang K., Cooper C., Yao H.,
Prusiner S.B., Hood L.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 9-232 FROM N.A.
RC TISSUE-BRAIN;
RX MEDLINE=93250789; PubMed=1363802;
RA Diedrich J.F., Knopman D.S., List J.F., Olson K., Frey W.H.,
Emory C.R., Sung J.H., Haase A.T.;
RT "Deletion in the prion protein gene in a demented patient.";
RL Hum. Mol. Genet. 1:443-444(1992).
RN [4]
RP SEQUENCE OF 41-85 FROM N.A.
RX MEDLINE=96090306; PubMed=7485229;
RA Perry R.T., Go R.C., Harrell L.E., Acton R.T.;
RT "SSCP analysis and sequencing of the human prion protein gene (PRNP) detects two different 24 bp deletions in an atypical Alzheimer's disease family.";
RL Am. J. Med. Genet. 60:12-18(1995).
DR EMBL: X83416; CAA58442.1; -;
DR EMBL: U29185; AAC78725.1; -;
DR EMBL: W81929; AAB59442.1; -;
DR EMBL: S80743; AAB50649.2; -;
DR EMBL: S80732; AAB50648.2; -;
DR HSP: P04156; IQLZ.
DR InterPro: IPR000817; Prion.
DR Pfam: PF00377; prion; 1.
DR PRINTS: PR00341; PRION.
DR SMART: SM00157; PRP; 1.
DR PROSITE: PS00291; PRION_1; 1.
DR PROSITE: PS00706; PRION_2; 1.
KW Prion.
SQ SEQUENCE 245 AA; 26884 MW; 6BF26E0FA3F061AD CRC64;

Query Match 95.2%; Score 1366; DB 4; Length 245;
Best Local Similarity 96.8%; Pred. No. 6.3e-118;
Matches 245; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 MANLGCMVLVLFVATWSDGLCKRKPXPGGNTGSGRYPGQSPGNGRYPPOGSGWGQP 60
DB 1 MANLGCMVLVLFVATWSDGLCKRKPXPGGNTGSGRYPGQSPGNGRYPPOG-----53
QY 61 HGGGQPHGGGQPHGGGQPHGGGQPHGGGQPHGGGQPHGGGQPHGGG 120

Db 240 VILLISFLFLVIG 253

RESULT 13

O46648 PRELIMINARY; PRT; 256 AA.

ID O46648; AC O46648; DT 01-JUN-1998 (TReMBLrel. 06, Created)

DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)

DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)

DE Prion protein variant 240S (Prion protein precursor).

GN PRP OR PRNP.

OS Capra hircus (Goat),

OS Ovis aries (Sheep),

OS Ovibos moschatus (Muskox),

OS Ovis canadensis (Bighorn sheep), and

OS Ovis orientalis musimon (Mouflon).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

OC Bovidae; Caprinae; Capra.

ON NCBI_TaxID=9925, 9940, 37176, 37174, 9938;

OX [1]

RP SEQUENCE FROM N.A.

RC SPECIES=C.hircus;

RC Vaccari G., Morelli L., Antonucci G., Agrimi U.;

RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.

[2]

RP SEQUENCE FROM N.A.

RC SPECIES=Sheep;

RC Cheung F., Goldmann W., Ford P., Hunter N.;

RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.

[3]

RP SEQUENCE FROM N.A.

RC SPECIES=Sheep;

RC MEDLINE=90207218; PubMed=1969635;

RX Goldmann W., Hunt N., Foster J.D., Salbaum J.M., Beyreuther K.,

RA Hope J.;

RT "Two alleles of a neural protein gene linked to scrapie in sheep.";

RT Proc. Natl. Acad. Sci. U.S.A. 87:2476-2480(1990).

[4]

RP SEQUENCE FROM N.A.

RC SPECIES=Sheep; TISSUE=BRAIN;

RC Lee I.Y., Westaway D., Smit A.F., Wang K., Cooper C., Yao H.,

RA Prusiner S.B., Hood L.;

RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.

[5]

RP SEQUENCE FROM N.A.

RC SPECIES=O.moschatus; TISSUE=PBL;

RC Schatzl H.M., Wopfinger F., Weidenhofer G., Gilch S.;

RT "Analysis of 27 mammalian and 9 avian prps reveals high conservation

RT of non-structural regions of the prion protein.";

RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.

[6]

RP SEQUENCE FROM N.A.

RC SPECIES=O.canadensis; TISSUE=BRAIN;

RC O'Rourke K.I., Spraker T.R., Wild M.A., Miller M.W.;

RT "Prp gene sequence for big horn sheep (Ovis canadensis).";

RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.

[7]

RP SEQUENCE FROM N.A.

RC SPECIES=O.orientalis musimon;

RC Seo S., Hara K., Kubosaki A., Nasu Y., Nishimura T., Saeki K.,

RA Matsumoto Y., Endo H., Onodera T.;

RT "Comparative analysis of the prion protein ORF nucleotide sequences

RT from two wild ruminants, mouflon and golden takin.;"

RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF496136; AAL93183.1;

DR EMBL; AJ223072; CAAL1073.1;

DR EMBL; U67922; AAC78726.1;

DR EMBL; AF117320; AADI9991.1;

DR EMBL; AF117316; AADI9987.1;

DR EMBL; AF156334; AAD48030.1;

Best Local Similarity 89.4%; Pred. No. 4.5e-113;			
Matches 227; Conservative 20; Mismatches 6; Indels 1; Gaps 1;			
QY	1	MANLGCWMLVLFVATWSDLGLCKKRPKPGGNTGSRYPGQSGGNRYPPQGGGGWQ	59
Db	1	MAHLGYWMLLLFVATWSVGLCKKRPKPGGNTGSRYPGQSGGNRYPPQGGGGWQ	60
QY	60	PHGGGGQPHGGGGQPHGGGGQPHGGGGQTHSQWNKPSKPKTNMKHMAGAAAAG	119
Db	61	PHGGGGQPHGGGGQPHGGGGQPHGGGGQTHSQWNKPSKPKTNMKHMAGAAAAG	120
QY	120	AVVGLGGYMLGSMSRPIIHFGSDYEDRYRENMHRYPNQVYRPMDEYSNQNNFVHDC	179
Db	121	AVVGLGGYMLGSMSRPLIHFGNDYEDRYRENMYRYPNQVYRPVDOYSNQNSFVHDC	180
QY	180	VNITIKOHTVTTTKGENTFTDVKMMRVVEQMCITQYERESQAYQYRGSSMVLFSPP	239
Db	181	VNITVKOHTVTTTKGENTFTDVKMMRVVEQMCITQYERESQAYQYRGSSMVLFSPP	240
QY	240	VILLISFLIFLIVG 253	
Db	241	VILLISFLIFLIVG 254	

Search completed: January 28, 2003, 14:18:16
Job time : 13.7848 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 13:58:08 ; Search time 14.0978 Seconds
(without alignments)
2391.326 Million cell updates/sec

Title: US-09-904-987-3

Perfect score: 1435

Sequence: 1 MANLGCWMLVLFATWSDLG.....LFSSPPVILLISFLIFLIVG 253

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*

1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1435	100.0	253	17 AAR86715	Human prion protei
2	1435	100.0	253	19 AAW69660	Human prion protei
3	1435	100.0	253	20 AAW85901	Human prion protei
4	1435	100.0	253	21 AAB15035	Human prion protei
5	1435	100.0	253	21 AAB06272	Human PrP pr
6	1435	100.0	253	21 AAY81485	Human prion protei
7	1435	100.0	253	21 AAG58553	Human prion protei
8	1435	100.0	253	22 AAB82112	Human PrP. Homo s
9	1435	100.0	253	22 AAB61770	Human prion protei
10	1435	100.0	253	23 ABP51787	Human prion protei

11	1435	100.0	253	23 AAE15603	Human PrP protein.
12	1435	100.0	253	23 ABE04426	Human prion protei
13	1427	99.4	253	20 AAY07994	Human prion protei
14	1425	99.3	253	22 AAB72338	Human prion protei
15	1423	99.2	253	23 AAB79575	Human prion protei
16	1422	99.1	253	23 AAB72341	Gorilla prion prot
17	1417	98.7	253	22 AAB72339	Chimpanzee prion p
18	1417	98.7	253	22 AAB72345	Gibbon prion prote
19	1417	98.7	253	22 AAB72356	Siayang prion prot
20	1404	97.8	253	22 AAB72340	Orangutan prion pr
21	1399	97.5	253	22 AAB72353	Guereza prion prot
22	1395	97.2	253	22 AAB72355	Prion protein cell
23	1392	97.0	253	22 AAB72350	Marmoset prion pro
24	1391	96.9	253	22 AAB72344	Rhesus monkey prio
25	1391	96.9	253	22 AAB72346	Prion protein cell
26	1391	96.9	253	22 AAB72347	Prion protein cell
27	1391	96.9	253	22 AAB72348	Prion protein cell
28	1391	96.9	253	22 AAB72349	Prion protein cell
29	1391	96.9	253	22 AAB72351	Hamadryas prion pr
30	1382	96.3	253	22 AAB72354	Capuchin prion pro
31	1376.5	95.9	260	22 AAB72343	Monkey prion prote
32	1332.5	92.9	264	22 AAB72361	Cow prion protein
33	1331.5	92.8	264	22 AAB82113	Bovine PrP. Bos t
34	1331.5	92.8	264	22 AAB61769	Bovine prion prote
35	1331.5	92.8	264	23 AAE15604	Bovine PrP protein
36	1329	92.6	263	17 AAR86716	Bovine prion prote
37	1329	92.6	263	19 AAW69661	Bovine prion prote
38	1329	92.6	263	20 AAW85902	Bovine prion prote
39	1329	92.6	263	22 AAG65854	Bovine prion prote
40	1329	92.6	263	23 ABP51788	Bovine prion prote
41	1326.5	92.4	264	23 ABB04424	Bovine prion prote
42	1323	92.2	265	23 AAM50889	Bovine prion prote
43	1322.5	92.2	284	20 AAY07995	Bovine prion prote
44	1321	92.1	255	17 AAR86717	Sheep prion protei
45	1321	92.1	255	19 AAW69662	Sheep prion protei

ALIGNMENTS

RESULT 1
AAR86715
ID AAR86715 standard; protein; 253 AA.
XX
AC AAR86715;
XX 15-OCT-1996 (first entry)
DT Human prion protein, HuPrP.
XX
DE
XX
KW Chimeric gene; chimeric prion; transgenic animal; diagnosis;
KW spongiform encephalopathy; PrP; central nervous system; CNS;
KW Creutzfeld-Jakob disease; CJD; BSE.
XX
XX Homo sapiens.
OS
PN W09531466-A1.
XX
XX 23-NOV-1995.
XX
XX 10-APR-1995; 95WO-US04426.
XX
XX 13-MAY-1994; 94US-0242188.
XX (REGC) UNIV CALIFORNIA.
XX Prusiner SB, Scott MR, Telling G;
XX WPT; 1996-010868/01.
XX Chimeric prion protein gene - for formation of a transgenic animal
XX susceptible to prion infection by prion(s) normally specific for a
XX different species

KW Human; PrP; prion; Creutzfeldt-Jakob disease; CJD; neuroprotective;
 KW vaccine; beta-form PrP; Kuru; transmissible mink encephalopathy;
 XX bovine spongiform encephalopathy; BSE.
 XX Homo sapiens.
 XX WO200026238-A2.
 XX 11-MAY-2000.
 XX 02-NOV-1999; 99WO-GB03617.
 XX 04-NOV-1998; 98GB-0024091.
 XX 18-MAR-1999; 99GB-0006217.
 XX (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
 XX Collinge J, Clarke AR, Jackson GS;
 XX WPI; 2000-365570/31.
 XX N-PSDB; AAA57174.
 XX Novel method of producing beta-form prion proteins, related antibodies,
 XX and binding agents useful in treatment and diagnosis of prion diseases
 XX .
 XX Disclosure; Fig 1; 115pp; English.
 XX The present sequence is the human prion protein encoded by the PrP
 XX gene. A beta-form of the protein has been produced which has more
 XX beta-sheet than alpha-helix structure, can exist as a monomer and
 XX can retain solubility in an aqueous solution in the absence of a
 XX denaturant. The beta-form PrP, a beta-form binding agent or aggregate can
 XX be used in the preparation or manufacture of a composition for the
 XX prevention, treatment and/or diagnosis of a prion disease, e.g. Kuru,
 XX Creutzfeldt-Jakob disease (CJD), transmissible mink encephalopathy,
 XX chronic wasting disease of mule deer and elk, and bovine spongiform
 XX encephalopathy (BSE). By measuring the levels of beta-form in a sample,
 XX agents capable of preventing, reducing and/or reversing the conversion of
 XX a PrP to a beta-form can be identified. The beta-form of the PrP or a
 XX non-fibrillar aggregate can be used as a vaccine against a prion disease.
 XX The beta-form can also be used to diagnose a predisposition to or the
 XX presence of a prion disease by monitoring aggregation.
 XX SQ Sequence 253 AA;
 Query Match 100.0%; Score 1435; DB 21; Length 253;
 Best Local Similarity 100.0%; Pred. No. 2.4e-135;
 Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MANLGCWMLVLFVATWSDGLCKKRPKPGWNTGSGRYPGQSPGNGRYPPOGGGGWGP 60
 Db 1 MANLGCWMLVLFVATWSDGLCKKRPKPGWNTGSGRYPGQSPGNGRYPPOGGGGWGP 60
 QY 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPKTNMKHMAGAAAAGA 120
 Db 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPKTNMKHMAGAAAAGA 120
 QY 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNFFVHDCV 180
 Db 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNFFVHDCV 180
 QY 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNFFVHDCV 180
 Db 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNFFVHDCV 180
 QY 181 NITIKQHTVTTTKGFNTETDVKMERVVEQMCITQYERESQAYYQRGSSWLVFSSPPV 240
 Db 181 NITIKQHTVTTTKGFNTETDVKMERVVEQMCITQYERESQAYYQRGSSWLVFSSPPV 240
 QY 241 ILLISFLIFLIVG 253
 Db 241 ILLISFLIFLIVG 253
 QY 181 NITIKQHTVTTTKGFNTETDVKMERVVEQMCITQYERESQAYYQRGSSWLVFSSPPV 240
 Db 181 NITIKQHTVTTTKGFNTETDVKMERVVEQMCITQYERESQAYYQRGSSWLVFSSPPV 240
 QY 241 ILLISFLIFLIVG 253
 Db 241 ILLISFLIFLIVG 253
 RESULT 6
 AAY81485

ID AAY81485 standard; Protein; 253 AA.
 XX AC AAY81485;
 XX DT 03-JUL-2000 (first entry)
 XX DE Human prion protein.
 XX KW Prion protein; human; antibody; monoclonal; prion-related disease;
 XX KW Creutzfeldt-Jakob disease; CJD; diagnosis.
 XX OS Homo sapiens.
 XX PN JP2000060551-A.
 XX PD 29-FEB-2000.
 XX PF 13-AUG-1998; 98JP-0241145.
 XX PR 13-AUG-1998; 98JP-0241145.
 XX PA (SEK) SEIKAGAKU KOGYO CO LTD.
 XX DR WPI; 2000-249674/22.
 XX Anti-prion antibody specific for a particular amino acid sequence -
 XX useful in diagnosis and research of prion related diseases
 XX PS Disclosure; Page 8; 9pp; Japanese.
 XX The invention relates to an anti-prion protein antibody, preferably a
 XX mouse-derived monoclonal antibody, which binds residues 96-114 of human
 XX prion protein (peptide sequence given in AAY81484). The anti-prion
 XX antibody is useful as a research tool and for the diagnosis of
 XX prion-related diseases such as Creutzfeldt-Jakob disease (CJD). The
 XX antibody is highly specific for human prion protein, and is relatively
 XX inexpensive to manufacture. This sequence represents full-length human
 XX prion protein.
 XX SQ Sequence 253 AA;
 Query Match 100.0%; Score 1435; DB 21; Length 253;
 Best Local Similarity 100.0%; Pred. No. 2.4e-135;
 Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MANLGCWMLVLFVATWSDGLCKKRPKPGWNTGSGRYPGQSPGNGRYPPOGGGGWGP 60
 Db 1 MANLGCWMLVLFVATWSDGLCKKRPKPGWNTGSGRYPGQSPGNGRYPPOGGGGWGP 60
 QY 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPKTNMKHMAGAAAAGA 120
 Db 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPKTNMKHMAGAAAAGA 120
 QY 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNFFVHDCV 180
 Db 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNFFVHDCV 180
 QY 181 NITIKQHTVTTTKGFNTETDVKMERVVEQMCITQYERESQAYYQRGSSWLVFSSPPV 240
 Db 181 NITIKQHTVTTTKGFNTETDVKMERVVEQMCITQYERESQAYYQRGSSWLVFSSPPV 240
 QY 241 ILLISFLIFLIVG 253
 Db 241 ILLISFLIFLIVG 253
 RESULT 7
 AAG65853
 ID AAG65853 standard; protein; 253 AA.
 XX AC AAG65853;
 XX DT 11-FEB-2002 (first entry)

```

XX DE Human prion protein (PrP) sequence.
XX DE
XX KW PrP; prion protein; Creutzfeldt-Jakob disease; familial insomnia; PrP-Sc;
XX KW scrapie; Gerstmann-Strassler-Scheinker disease.
XX OS Homo sapiens.
XX PN US6290954-B1.
XX PD 18-SEP-2001.
XX PF 06-MAR-1998; 98US-0036579.
XX PR 13-SEP-1996; 96US-0713939.
XX PR 14-SEP-1995; 95US-0528104.
XX PA (SCRI ) SCRIPPS RES INST.
XX PI Prusiner SB, Williamson RA, Burton DR;
XX WPI; 2001-637939/73.
XX
XX PT Detecting a scrapie isoform of the prion protein (PrP-Sc) in a source,
XX PT particularly useful for detecting e.g. Creutzfeldt-Jakob disease or
XX PT Gerstmann-Strassler-Scheinker disease, by contacting the source with
XX PT PrP-Sc antibodies -
XX PS Disclosure; Fig 2; 58pp; English.
XX
XX CC The invention provides a method for detecting a scrapie isoform of the
XX CC prion protein (PrP-Sc) in a source. The method involves contacting the
XX CC source suspected of containing native PrP-Sc with a diagnostic amount of
XX CC an antibody characterized by its ability to bind to native PrP-Sc in
XX CC situ. The method is useful for detecting PrP-Sc in a source, which is
XX CC particularly useful for detecting Creutzfeldt-Jakob disease, fatal
XX CC familial insomnia or Gerstmann-Strassler-Scheinker disease. The present
XX CC sequence represents the human PrP sequence.
XX
XX SQ Sequence 253 AA;
XX
XX Query Match 100.0%; Score 1435; DB 22; Length 253;
XX Best Local Similarity 100.0%; Pred. No. 2.4e-135;
XX Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MANLGCMVLVLFVATWSDLGLCKKRPKPGGNTGSGRYPGQSPGNGRYPPOGGGGWGQP 60
DB 1 MANLGCMVLVLFVATWSDLGLCKKRPKPGGNTGSGRYPGQSPGNGRYPPOGGGGWGQP 60
QY 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGHGTHSQWNKPSKPTNNKHMAGAAAAGA 120
DB 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGHGTHSQWNKPSKPTNNKHMAGAAAAGA 120
QY 121 VVGLGGLGMLGSAMSRPIIHFGSDYEDRYRENHMYRNQVYRPMDEYSNQNNFVHDCV 180
DB 121 VVGLGGLGMLGSAMSRPIIHFGSDYEDRYRENHMYRNQVYRPMDEYSNQNNFVHDCV 180
QY 181 NITIKOHTVTTTKGENFTETDVKMVRVVEQMCITQYERESQAYYQRGSSMVLFSPPV 240
DB 181 NITIKOHTVTTTKGENFTETDVKMVRVVEQMCITQYERESQAYYQRGSSMVLFSPPV 240
QY 241 ILLISFLIFLIVG 253
DB 241 ILLISFLIFLIVG 253
XX
XX RESULT 8
XX ID AAB82112
XX AC AAB82112 standard; Protein; 253 AA.
XX DT 29-JUN-2001 (first entry)

```

```

XX DE Human Prp.
XX DE
XX KW Human; Prp; cerebroprotective; PrP conversion inhibitor; prion protein;
XX KW transmissible spongiform encephalopathy; tSE; neurodegenerative disease;
XX KW protease-sensitive prion protein; PrPsen;
XX KW protease-resistant prion protein; PrPres.
XX OS Homo sapiens.
XX PN US6211149-B1.
XX PD 03-APR-2001.
XX PF 03-AUG-1998; 98US-0128450.
XX PR 03-AUG-1998; 98US-0128450.
XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX PI Chesebro BW, Caughey BW, Chabry J, Priola S;
XX WPI; 2001-315407/33.
XX
XX PT New peptide comprises hamster prion protein fragment that inhibits
XX PT conversion of prion protein from protease-sensitive to
XX PT protease-resistant form, useful for diagnosis and treatment of
XX PT spongiform encephalopathies -
XX PS Disclosure; Column 33-34; 31pp; English.
XX
XX CC Transmissible spongiform encephalopathies (TSE) are fatal
XX CC neurodegenerative diseases. These diseases are characterised by the
XX CC formation and accumulation, in the brain, of an abnormal proteinase K
XX CC resistant isoform (PrPres) of a normal protease-sensitive host-encoded
XX CC prion protein (PrPsen). The present invention relates to peptides
XX CC comprising a hamster, human or murine prion protein (PrP) fragment which
XX CC specifically inhibit the conversion of protease-sensitive prion protein
XX CC (PrPsen) to protease-resistant prion protein (PrPres). The present
XX CC sequence is the protein sequence for human PrP. The peptides of the
XX CC present invention are useful for diagnosis and treatment of TSE
XX CC diseases.
XX
XX SQ Sequence 253 AA;
XX
XX Query Match 100.0%; Score 1435; DB 22; Length 253;
XX Best Local Similarity 100.0%; Pred. No. 2.4e-135;
XX Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MANLGCMVLVLFVATWSDLGLCKKRPKPGGNTGSGRYPGQSPGNGRYPPOGGGGWGQP 60
DB 1 MANLGCMVLVLFVATWSDLGLCKKRPKPGGNTGSGRYPGQSPGNGRYPPOGGGGWGQP 60
QY 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGHGTHSQWNKPSKPTNNKHMAGAAAAGA 120
DB 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGHGTHSQWNKPSKPTNNKHMAGAAAAGA 120
QY 121 VVGLGGLGMLGSAMSRPIIHFGSDYEDRYRENHMYRNQVYRPMDEYSNQNNFVHDCV 180
DB 121 VVGLGGLGMLGSAMSRPIIHFGSDYEDRYRENHMYRNQVYRPMDEYSNQNNFVHDCV 180
QY 181 NITIKOHTVTTTKGENFTETDVKMVRVVEQMCITQYERESQAYYQRGSSMVLFSPPV 240
DB 181 NITIKOHTVTTTKGENFTETDVKMVRVVEQMCITQYERESQAYYQRGSSMVLFSPPV 240
QY 241 ILLISFLIFLIVG 253
DB 241 ILLISFLIFLIVG 253
XX
XX RESULT 9
XX ID AAB61770
XX ID AAB61770 standard; protein; 253 AA.

```

Db	181	NITKQHTVTTTKGNFTETDVKMWERVVEQMCITQYERESQAYQRGSMVLFSSPPV	240
QY	241	ILLISFLIFLIVG	253
Db	241	ILLISFLIFLIVG	253
RESULT 10			
ABP51787			
ID	ABP51787	standard; protein; 253 AA.	
XX			
AC	ABP51787;		
XX			
DT	03-OCT-2002	(first entry)	
XX			
DE	Human prion protein (PrP)	SEQ ID NO:2.	
XX			
KW	Prion protein; PrP; scrapie; PrPSc; prion disease; immunoassay;		
KW	detection.		
XX			
OS	Homo sapiens.		
XX			
PN	US6372214-B1.		
XX			
PD	16-APR-2002.		
XX			
PF	13-APR-2000; 2000US-0550374.		
XX			
PR	13-SEP-1996; 96US-0713939.		
PR	06-MAR-1998; 98US-0036579.		
PR	14-SEP-1995; 95US-0528104.		
XX			
PA	(REGC) UNIV CALIFORNIA.		
PA	(SCRI) SCRIPPS RES INST.		
XX			
PI	Prusiner SB, Williamson RA, Burton DR;		
XX			
DR	WPI; 2002-433675/46.		
XX			
PT	Immunoassays for detecting scrapie isoforms of prion protein (PrPSc)		
PT	and for purifying PrPSc from samples, useful e.g. in diagnosing PrPSc		
PT	disease and testing pharmaceuticals for contamination -		
XX			
PS	Disclosure; Fig 2: 58pp; English.		
XX			
CC	The present invention describes methods for detecting scrapie isoforms		
CC	of prion protein (PrPSc) infection in dead animals, purifying materials		
CC	suspected of containing PrPSc proteins and treating materials, using		
CC	antibodies specific for PrPSc. Also described: (1) method of determining		
CC	PrPSc infection in a dead animal, comprising: (a) extracting tissue from		
CC	an animal that has died; (b) contacting the tissue with an antibody		
CC	characterised by its ability to bind to native PrPSc in situ (the		
CC	antibody binds to a form of PrPSc specific to the animal that has died);		
CC	and (c) determining if the antibody has bound to PrPSc (the presence of		
CC	PrPSc in the tissue is indicative of PrPSc infection); (2) a method of		
CC	purifying a material suspected of containing a PrPSc protein, comprising:		
CC	(a) contacting the material with an antibody (characterized by its		
CC	ability to bind native PrPSc in situ) which is bound to a support		
CC	surface; and (b) removing material not bound to the antibody; (3) a		
CC	method of treating a material, comprising applying (to the material) an		
CC	antibody that binds native PrPSc in situ. The methods are used for		
CC	diagnosing and detecting prion disease (scrapie) in dead animal tissue		
CC	(i.e. immunoassays), for separating PrPSc proteins from biological		
CC	samples (i.e. immunopurification) and for treating materials. The present		
CC	sequence represents the human prion protein (PrP) which is given in the		
CC	embodiment of the present invention.		
XX			
Sequence	253 AA;		
Query Match	100.0%; Score 1435; DB 23; Length 253;		
Best Local Similarity	100.0%; Pred. No. 2.4e-135;		
Matches 253; Conservative	0; Mismatches 0; Indels 0; Gaps 0;		

QY 1 MANLGCMMLVLFVATWSDLGLCKKRPKPGWNTGSGRYPGQSPGNGNRYPPQGGGGWGQ 60
Db 1 MANLGCMMLVLFVATWSDLGLCKKRPKPGWNTGSGRYPGQSPGNGNRYPPQGGGGWGQ 60
QY 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQNNKPSKPTNNKHMAAAGAA 120
Db 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQNNKPSKPTNNKHMAAAGAA 120
QY 121 VVGLGGLYMLGSAMSRPIIHFGSDYEDRYRENHMYRNQVYRPMDEYSNQNNEFVHDCV 180
Db 121 VVGLGGLYMLGSAMSRPIIHFGSDYEDRYRENHMYRNQVYRPMDEYSNQNNEFVHDCV 180
QY 181 NITIKOHTVTTTNGENFTETDVKMERVVEQMCITQYERESQAYYQRGSSMVLFSPPV 240
Db 181 NITIKOHTVTTTNGENFTETDVKMERVVEQMCITQYERESQAYYQRGSSMVLFSPPV 240
RESULT 11
AAE15603
ID AAE15603 standard; Protein; 253 AA.
AC AAE15603;
DT 12-MAR-2002 (first entry)
DE Human PrP protein.
KW Protease resistant prion protein; PrPres; Creutzfeldt-Jakob disease;
KW protease sensitive prion protein; PrPsen; therapy; neuroprotectant;
KW transmissible spongiform encephalopathy; fatal familial insomnia;
KW Gerstmann-Straussler-Scheinker syndrome; scrapie; kuru; human.
OS Homo sapiens.
XX US2001041790-A1.
PN 15-NOV-2001.
XX 30-MAR-2001; 2001US-0823494.
XX 12-MAY-1998; 98US-085160P.
PR 03-AUG-1998; 98US-0128450.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX Chesebro BW, Caughey BW, Chabry J, Priola S;
XX WPI; 2002-065984/09.
XX Peptides which can inhibit conversion of prion proteins from protease
PT sensitive to protease resistant forms, are useful in treating
PT transmissible spongiform encephalopathies -
XX Claim 45; Page 18-19; 33pp; English.
XX The present invention relates to peptides comprising a peptide region of
CC prion protein (PrP) and which are capable of inhibiting conversion of
CC protease sensitive prion protein (PrPsen) to protease resistant prion
CC protein (PrPres). The peptides are useful for inhibiting formation of
CC protease resistant prion proteins, such as those associated with
CC transmissible spongiform encephalopathies, e.g., Creutzfeldt-Jakob
CC disease, kuru, Gerstmann-Straussler-Scheinker syndrome, fatal familial
CC insomnia or scrapie. The peptides can be used as diagnostic agents, e.g.,
CC to detect the presence of PrPres in body fluids such as blood. They may
CC also be used to treat or prevent diseases such as the above. They may
CC also be used to design analogues, derivatives or mimetics for use as
CC inhibitors of conversion of PrPsen to PrPres. The present sequence is
CC human PrP protein.
XX

SQ Sequence 253 AA;
Query Match 100.0%; Score 1435; DB 23; Length 253;
Best Local Similarity 100.0%; Pred. No. 2,4e-135;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MANLGCMMLVLFVATWSDLGLCKKRPKPGWNTGSGRYPGQSPGNGNRYPPQGGGGWGQ 60
Db 1 MANLGCMMLVLFVATWSDLGLCKKRPKPGWNTGSGRYPGQSPGNGNRYPPQGGGGWGQ 60
QY 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQNNKPSKPTNNKHMAAAGAA 120
Db 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQNNKPSKPTNNKHMAAAGAA 120
QY 121 VVGLGGLYMLGSAMSRPIIHFGSDYEDRYRENHMYRNQVYRPMDEYSNQNNEFVHDCV 180
Db 121 VVGLGGLYMLGSAMSRPIIHFGSDYEDRYRENHMYRNQVYRPMDEYSNQNNEFVHDCV 180
QY 181 NITIKOHTVTTTNGENFTETDVKMERVVEQMCITQYERESQAYYQRGSSMVLFSPPV 240
Db 181 NITIKOHTVTTTNGENFTETDVKMERVVEQMCITQYERESQAYYQRGSSMVLFSPPV 240
QY 241 ILLISFLIFLI 253
Db 241 ILLISFLIFLI 253
RESULT 12
ABB04426
ID ABB04426 standard; Protein; 253 AA.
XX AC ABB04426;
XX DT 04-MAR-2002 (first entry)
XX DE Human prion protein PrP.
XX KW Human; prion protein; PrP; antiviral; HIV; prion disease; kuru; virucide;
KW antibacterial; neuroprotective; anti-HIV; Creutzfeldt-Jakob disease;
KW Gerstmann-Straussler-Scheinker disease; fatal familial insomnia;
KW bovine spongiform encephalitis; scrapie.
XX OS Homo sapiens.
XX WO200183747-A2.
XX PD 08-NOV-2001.
XX PF 30-APR-2001; 2001WO-FR01336.
XX PR 28-APR-2000; 2000FR-0005535.
XX PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
XX PI Leblanc P, Darlix J, Gabus-Darlix C;
XX DR WPI; 2002-049350/06.
XX New polypeptides, useful as antiviral agents, comprise their prion
PT proteins able to bind nucleic acid, nucleocapsid proteins, and ligands
PT for use as antiprion agents -
XX Disclosure; Fig 9; 80pp; French.
XX The present invention relates to normal (PrPc) or abnormal (PrPsc) human
CC or animal prion proteins which are able to bind to DNA or RNA,
CC particularly of viral, especially retroviral, origin and to nucleocapsid
CC proteins (NCP) of human or animal retroviruses. These can be used as
CC antiviral agents, particularly against human immune deficiency virus
CC (HIV), and in the treatment of prion diseases including Creutzfeldt-Jakob
CC disease, Gerstmann-Straussler-Scheinker disease, fatal familial
CC insomnia, kuru, bovine spongiform encephalitis and scrapie. The present
CC sequence is the human PrP protein sequence.
XX

QY 241 ILLISFLIFLIVG 253
|||||
Db 241 ILLISFLIFLIVG 253

RESULT 2

US-09-943-906-2
; Sequence 2, Application US/09943906
; Patent No. US20020150571A1
; GENERAL INFORMATION:
; APPLICANT: Prusiner, Stanley B.
; William, R. Anthony
; Burton, Dennis R.
; TITLE OF INVENTION: ANTIBODIES SPECIFIC FOR NATIVE PrP
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 2200 Sand Hill Road
; CITY: Menlo Park
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA: US/09/943,906
; FILING DATE: 30-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/550,374
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Bozicevic, Karl
; REGISTRATION NUMBER: 28,807
; REFERENCE/DOCKET NUMBER: 06510/059001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-854-5277
; TELEFAX: 415-854-0875
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-943-906-2

Query Match 100.0%; Score 1435; DB 10; Length 253;
Best Local Similarity 100.0%; Pred. No. 8.5e-117;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MANLGCMVLVLFVATWSDGLCKKRPKPGGNTGSGRYPGQSGPGNRYPPQGGGGWGQP 60
Db 1 MANLGCMVLVLFVATWSDGLCKKRPKPGGNTGSGRYPGQSGPGNRYPPQGGGGWGQP 60
QY 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPKTNMKHMAGAAAAA 120
Db 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPKTNMKHMAGAAAAA 120
QY 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNMFVHDCV 180
Db 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNMFVHDCV 180
QY 181 NITIKQHTVTTTNGENFTETDVKMERVVEQMCITQYERESQAYYQRGSSMVLFSPPV 240
Db 181 NITIKQHTVTTTNGENFTETDVKMERVVEQMCITQYERESQAYYQRGSSMVLFSPPV 240

QY 241 ILLISFLIFLIVG 253
|||||
Db 241 ILLISFLIFLIVG 253

RESULT 3

US-09-919-172-57
; Sequence 57, Application US/09919172
; Patent No. US20020119463A1
; GENERAL INFORMATION:
; APPLICANT: Paris, Mary
; APPLICANT: Turner, Christopher M.
; TITLE OF INVENTION: PROSTATE CANCER MARKERS
; FILE REFERENCE: PA-0036 US
; CURRENT APPLICATION NUMBER: US/09/919,172
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/222,469
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PERL Program
; SEQ ID NO 57
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020119463A1 1256895CD1
US-09-919-172-57

Query Match 99.7%; Score 1431; DB 10; Length 253;
Best Local Similarity 99.6%; Pred. No. 1.9e-116;
Matches 252; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MANLGCMVLVLFVATWSDGLCKKRPKPGGNTGSGRYPGQSGPGNRYPPQGGGGWGQP 60
Db 1 MANLGCMVLVLFVATWSDGLCKKRPKPGGNTGSGRYPGQSGPGNRYPPQGGGGWGQP 60
QY 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPKTNMKHMAGAAAAA 120
Db 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPKTNMKHMAGAAAAA 120
QY 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNMFVHDCV 180
Db 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNMFVHDCV 180
QY 181 NITIKQHTVTTTNGENFTETDVKMERVVEQMCITQYERESQAYYQRGSSMVLFSPPV 240
Db 181 NITIKQHTVTTTNGENFTETDVKMERVVEQMCITQYERESQAYYQRGSSMVLFSPPV 240
QY 241 ILLISFLIFLIVG 253
|||||
Db 241 ILLISFLIFLIVG 253

RESULT 4

US-09-943-906-3
; Sequence 3, Application US/09943906
; Patent No. US20020150571A1
; GENERAL INFORMATION:
; APPLICANT: Prusiner, Stanley B.
; William, R. Anthony
; Burton, Dennis R.
; TITLE OF INVENTION: ANTIBODIES SPECIFIC FOR NATIVE PrP
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 2200 Sand Hill Road
; CITY: Menlo Park
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette

```

US-10-109-551-2
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/09/943,906
  FILING DATE: 30-Aug-2001
  CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 09/550,374
  FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
  NAME: Bozicevic, Karl
  REGISTRATION NUMBER: 28,807
  REFERENCE/DOCKET NUMBER: 06510/059001
TELECOMMUNICATION INFORMATION:
  TELEPHONE: 415-854-5277
  TELEFAX: 415-854-0875
  TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 3:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 263 amino acids
    TYPE: amino acid
    STRANDEDNESS: single
    TOPOLOGY: linear
  MOLECULE TYPE: peptide
  SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-943-906-3
Query Match          92.6%; Score 1329; DB 10; Length 263;
Best Local Similarity 87.7%; Pred. No. 1.2e-107;
Matches 228; Conservative 19; Mismatches 5; Indels 8; Gaps 1;
QY 2 ANLCWMLVLFVATWSDLGLCKKRPKPGGNTGSGRYPGGSPGNGRYPPQGGGQGP 61
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 4 SHIGSMILVLFVAMWSDVGLCKKRPKPGGNTGSGRYPGGSPGNGRYPPQGGGQGP 63
QY 62 GGGWQPHGGGQGP 61
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 64 GGGWQPHGGGQGP 61
QY 114 GAAAGAVVGLGGLGYSMSRPLIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQN 173
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 124 GAAAGAVVGLGGLGYSMSRPLIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQN 183
QY 174 NFVHDCVNITIKQHTVTTTGTGNTFTDVKMERVVEQMCITQYERESQAYYQGS 233
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 184 NFVHDCVNITIKQHTVTTTGTGNTFTDVKMERVVEQMCITQYERESQAYYQGS 243
QY 234 LFSPPVILLISFLIFLVG 253
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 244 LFSPPVILLISFLIFLVG 263
RESULT 5
US-10-109-551-2
Sequence 2, Application US/10109551
Publication No. US20020194635A1
GENERAL INFORMATION:
  APPLICANT: DUNNE, PATRICK W.
  APPLICANT: PIEDRAHITA, JORGE
  TITLE OF INVENTION: TRANSGENIC ANIMALS RESISTANT TO TRANSMISSIBLE
  FILE REFERENCE: TAMK:207US
  CURRENT APPLICATION NUMBER: US/10/109,551
  PRIOR FILING DATE: 2002-03-28
  PRIOR APPLICATION NUMBER: 60/280,549
  PRIOR FILING DATE: 2001-03-30
  NUMBER OF SEQ ID NOS: 10
  SOFTWARE: PatentIn Ver. 2.1
  SEQ ID NO 2
  LENGTH: 256
  TYPE: PRT
  ORGANISM: Bos taurus
US-10-109-551-2
Query Match          92.4%; Score 1326.5; DB 9; Length 256;
Best Local Similarity 91.3%; Pred. No. 2e-107;
Matches 232; Conservative 16; Mismatches 3; Indels 3; Gaps
QY 2 ANLCWMLVLFVATWSDLGLCKKRPKPGGNTGSGRYPGGSPGNGRYPPQGGGQGP 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 4 SHIGSMILVLFVAMWSDVGLCKKRPKPGGNTGSGRYPGGSPGNGRYPPQGGGQGP 63
QY 61 HGGWQPHGGGQGP 61
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 64 HGGWQPHGGGQGP 61
QY 120 AVVGLGGYMLGYSMSRPLIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQN 179
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 123 AVVGLGGYMLGYSMSRPLIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQN 182
QY 180 VNITIKQHTVTTTGTGNTFTDVKMERVVEQMCITQYERESQAYYQGS 233
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 183 VNITIKQHTVTTTGTGNTFTDVKMERVVEQMCITQYERESQAYYQGS 242
QY 240 VILLISFLIFLVG 253
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 243 VILLISFLIFLVG 256
RESULT 6
US-09-943-906-4
Sequence 4, Application US/09943906
Patent No. US20020150571A1
GENERAL INFORMATION:
  APPLICANT: Prusiner, Stanley B.
  APPLICANT: Williamson, R. Anthony
  APPLICANT: Burton, Dennis R.
  TITLE OF INVENTION: ANTIBODIES SPECIFIC FOR NATIVE PrP
  NUMBER OF SEQUENCES: 86
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: Fish & Richardson P.C.
    STREET: 2200 Sand Hill Road
    CITY: Menlo Park
    STATE: CA
    COUNTRY: U.S.A.
    ZIP: 94025
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Diskette
    COMPUTER: IBM Compatible
    OPERATING SYSTEM: DOS
    SOFTWARE: FastSeq Version 2.0
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/09/943,906
    FILING DATE: 30-Aug-2001
    CLASSIFICATION: <Unknown>
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: 09/550,374
    FILING DATE: <Unknown>
  ATTORNEY/AGENT INFORMATION:
    NAME: Bozicevic, Karl
    REGISTRATION NUMBER: 28,807
    REFERENCE/DOCKET NUMBER: 06510/059001
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: 415-854-5277
    TELEFAX: 415-854-0875
    TELEX: <Unknown>
  INFORMATION FOR SEQ ID NO: 4:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 255 amino acids
      TYPE: amino acid
      STRANDEDNESS: single
      TOPOLOGY: linear
    MOLECULE TYPE: peptide
    SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-943-906-4

```


QY 2 ANLGCMMLVLFVATWSDGLCKRKP-KGWNTGSRYPGQSPGNRYPPQGGGGWGP 60
Db 4 SHIGSWILVLFVAMWSDVGLCKRKP-KGWNTGSRYPGQSPGNRYPPQGGGGWGP 63
QY 61 HGGWCQPHGGGWGQPHGGGWGQPH-GGGWGGGGTHSQWNKPSKPKTNMKHMAGAAAG 119
Db 64 HGGWCQPHGGGWGQPHGGGWGQPHGGGWGQ-GGTHSQWNKPSKPKTNMKHMAGAAAG 122
QY 120 AVVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNFFVHDC 179
Db 123 AVVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNFFVHDC 182
QY 180 VNITKQHTVTTTNGENFTETDKMMRVVEQMCITQYRESQAYYQRGSSMVLFSPP 239
Db 183 VNITKQHTVTTTNGENFTETDKMMRVVEQMCITQYRESQAYYQRGSSMVLFSPP 242
QY 240 VILLISFLIFLIVG 253
Db 243 VILLISFLIFLIVG 256

RESULT 10

US-10-109-551-10
; Sequence 10, Application US/10109551
; Publication No. US20020194635A1
; GENERAL INFORMATION:
; APPLICANT: DUNNE, PATRICK W.
; APPLICANT: PIEDRAHITA, JORGE
; TITLE OF INVENTION: TRANSGENIC ANIMALS RESISTANT TO TRANSMISSIBLE
; FILE REFERENCE: TAMK:207US
; CURRENT APPLICATION NUMBER: US/10/109,551
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: 60/280,549
; PRIOR FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Cervus elaphus
US-10-109-551-10

Query Match 91.5%; Score 1312.5; DB 9; Length 256;
Best Local Similarity 90.2%; Pred. No. 3.2e-106;
Matches 229; Conservative 19; Mismatches 3; Indels 3; Gaps 3;

QY 2 ANLGCMMLVLFVATWSDGLCKRKP-KGWNTGSRYPGQSPGNRYPPQGGGGWGP 60
Db 4 SHIGSWILVLFVAMWSDVGLCKRKP-KGWNTGSRYPGQSPGNRYPPQGGGGWGP 63
QY 61 HGGWCQPHGGGWGQPHGGGWGQPH-GGGWGGGGTHSQWNKPSKPKTNMKHMAGAAAG 119
Db 64 HGGWCQPHGGGWGQPHGGGWGQPHGGGWGQ-GGTHSQWNKPSKPKTNMKHMAGAAAG 122
QY 120 AVVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNFFVHDC 179
Db 123 AVVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNFFVHDC 182
QY 180 VNITKQHTVTTTNGENFTETDKMMRVVEQMCITQYRESQAYYQRGSSMVLFSPP 239
Db 183 VNITKQHTVTTTNGENFTETDKMMRVVEQMCITQYRESQAYYQRGSSMVLFSPP 242
QY 240 VILLISFLIFLIVG 253
Db 243 VILLISFLIFLIVG 256

RESULT 11

US-10-106-574-5
; Sequence 5, Application US/10106574
; Patent No. US20020164335A1
; GENERAL INFORMATION:

; APPLICANT: Harris, David A.
; APPLICANT: Stewart, Richard S.
; TITLE OF INVENTION: Compositions and Methods for the Study and Diagnosis of Prion
; FILE REFERENCE: 09789280.0003
; CURRENT APPLICATION NUMBER: US/10/106,574
; CURRENT FILING DATE: 2002-03-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
US-10-106-574-5

Query Match 90.6%; Score 1299.5; DB 9; Length 254;
Best Local Similarity 89.4%; Pred. No. 4.2e-105;
Matches 228; Conservative 15; Mismatches 9; Indels 3; Gaps 2;

QY 1 MANLGCMMLVLFVATWSDGLCKRKP-KGWNTGSRYPGQSPGNRYPPQGGGGWGP 60
Db 1 MANLGCMMLVLFVATWSDGLCKRKP-KGWNTGSRYPGQSPGNRYPPQGGGGWGP 59
QY 61 HGGWCQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPKTNMKHMAGAAAG 120
Db 60 HGGWCQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPKTNMKHMAGAAAG 119
QY 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNFFVHDCV 180
Db 120 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNFFVHDCV 179
QY 181 NITIKQHTVTTTNGENFTETDKMMRVVEQMCITQYRESQAYY--ORGSSMVLFSPP 238
Db 180 NITIKQHTVTTTNGENFTETDKMMRVVEQMCITQYRESQAYY--ORGSSMVLFSPP 239
QY 239 PVILLISFLIFLIVG 253
Db 240 PVILLISFLIFLIVG 254

RESULT 12

US-09-943-906-1
; Sequence 1, Application US/09943906
; Patent No. US20020150571A1
; GENERAL INFORMATION:
; APPLICANT: Prusiner, Stanley B.
; APPLICANT: Williamson, R. Anthony
; APPLICANT: Burton, Dennis R.
; TITLE OF INVENTION: ANTIBODIES SPECIFIC FOR NATIVE PrP
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 2200 Sand Hill Road
; CITY: Menlo Park
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/943,906
; FILING DATE: 30-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/550,374
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Bozicevic, Karl
; REGISTRATION NUMBER: 28,807
; REFERENCE/DOCKET NUMBER: 06510/059001
; TELECOMMUNICATION INFORMATION:

```

, , TELEPHONE: 415-854-5277
, ,
, , TELEFAX: 415-854-0875
, ,
, , TELE: <Unknown>
, ,
, , INFORMATION FOR SEQ ID 1:
, ,
, , SEQUENCE CHARACTERISTICS:
, ,
, , LENGTH: 254 amino acids
, ,
, , TYPE: amino acid
, ,
, , STRANDEDNESS: single
, ,
, , TOPOLOGY: linear
, ,
, , MOLECULE TYPE: peptide
, ,
, , SEQUENCE DESCRIPTION: SEQ ID
, ,
US-09-943-906-1

```

Query Match 90.3%; Score 1296.5; DB 10; Length 254;
Best Local Similarity 89.0%; Pred. NO. 7.6e-105;
Matches 227; Conservative 16; Mismatches 9; Indels 3; Gaps 2;

Qy	1	MANLGCWMLVLFVATVSDSLGCKRKPKPGWNTPGGSRYPQGSGPGGNRYPPGGGGWGQP	60
Db	1	MANLGYLLALFVMTWTDVGLCKRKPKPGWNTPGGSRYPQGSGPGGNRYPPQ-GGTWGQP	59
Qy	61	HGGGSGP HGGGSGQP HGGGSGOP HGGGSGOGGTHSONNKPSPKPTNKKHAGAAAAAGA	120
Db	60	HGGGSGP HGGGSGQP HGGGSGOP HGGGSGOGGTHINQNNKPSKPTNLKHVAGAAAAAGA	119
Qy	121	VVGGLGYMLGSAMSRIPIHFGSDYEDRYRRENNHRYPNQVYRPMDEYSNQNNFVHDCV	180
Db	120	VVGGLGYMLGSAMSRLPIHFGNDWEDRYRRENNRYPNQVYRYPDQYSNQNNFVHDCV	179
Qy	181	NITIKQHTVTTTKGKNFTETDVKMMRVVEQMCITQYERESQAY--ORGSSWVLFSSP	238
Db	180	NITIKQHTVTTTKGKNFTETDVKMMRVVEQMCVTQYQESQAYYDGRSSSTVLFSSP	239
Qy	239	PVLLISFLIFLIVG	253
Db	240	PVLLISFLIFLIVG	254

```

RESULT 13
US-10-106-574-6
; Sequence 6, Application US/10106574
; Patent No. US20020164395A1
; GENERAL INFORMATION:
; APPLICANT: Harris, David A.
; APPLICANT: Stewart, Richard S.
; TITLE OF INVENTION: Compositions and
; FILE REFERENCE: 09789280.0003
; CURRENT APPLICATION NUMBER: US/10/
; CURRENT FILING DATE: 2002-03-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 254
; TYPE: prt
; ORGANISM: Murinae gen. sp.
US-10-106-574-6

```

Query Match 90.1%; Score 1293.5; DB 9; Length 254;
Best Local Similarity 89.0%; Pred. No. 1.4e-104;
Matches 227; Conservative 15; Mismatches 10; Indels 3; Gaps 2;

[illegible]

```

Qy   181 NITTKQHVTVTTTNGENFTETDVKKMERVVYEQMCITQYERESQAYY--QRGSSNMWLFSSP 238
      |||||
Db   180 NITIKQHTVTTTTKGENFTETDVKKMERVVYEQMCVTQYQESQAYYDGRSSSTVLFFSP 239
      |||||

Qy   239 PVIILLISFLIFLIVG 253
      |||||
Db   240 PVIILLISFLIFLIVG 254
      |||||

RESULT 14
US-10-106-574-7
; Sequence 7, Application US/10106574
; Patent No. US20020164335A1
; GENERAL INFORMATION:
; APPLICANT: Harris, David A.
; APPLICANT: Stewart, Richard S.
```

Query Match	89.7%	Score 1287.5	DB 9	Length 254
Best Local Similarity	88.2%	Pred. No. 4.5e-104		
Matches 225	Conservative 15	Mismatches 12	Indels 3	Gaps 2
Qy 1	MANLGCMVLVLVATWSDLG	LCKRPRPGGWN	TGGSRYPGGSPG	GNRYPRQGSGGWCQP 60
	:			
Db 1	MANLGWLLALFVMTDVG	LCKRPRPGGWN	TGGSRYPGGSPG	GNRYPPQ-GGTWQCP 59
	:			
Qy 61	HGGGKGQPHGGGWCQP	HGGGWCQGGTGH	SHQNNKPSR	PKTKNKHMACAAAAA 120
	:			
Db 60	HGGGKGQPHGGGWCQP	HGGGWCQGGTGH	SHQNNKPSR	PKTKNKHMGVAAVGA 119
	:			
Qy 121	VVGLGGYMLGSAMSRP	LIHFSGSYEDRY	YRENMHRYPNQV	YYPMDSEYSNQNNFVHDCV 180
	:			
Db 120	VVGLGGYMLGSVSRPM	IHFNGDWDERY	YRENMYRYPNQV	YYPMDSEYSNQNNFVHDCV 179
	:			
Qy 181	NITIKQHTVTTTTKGN	FTETDVKMMERVV	BQMCITQYERESQ	AY--QRGSSMVLFSPP 238
	:			
Db 180	NITIKQHTVTTTTKGN	FTETDVKMMERVV	BQMCITQYERESQ	AY--QRGSSMVLFSPP 239
	:			
Qy 239	PVILLISFLIFLIVG	253		
	:			
Db 240	PVILLISFLIFLIVG	254		
	:			

```

RESULT 15
US-10-106-574-8
; Sequence 8, Application US/10106574
; Patent No. US20020164335A1
; GENERAL INFORMATION:
; APPLICANT: Harris, David A.
; APPLICANT: Stewart, Richard S.
; TITLE OF INVENTION: Compositions and Methods for the Study and Diagnosis of Prion
; FILE REFERENCE: 09789280.0003
; CURRENT APPLICATION NUMBER: US/10/106,574
; CURRENT FILING DATE: 2002-03-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 254
; TYPE: prt
; ORGANISM: Murinae gen. sp.
US-10-106-574-8
Query Match      89.3%   Score 1281.5;   DB 9;   Length 254;

```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:10:57 : Search time 5.25212 Seconds
(without alignments)
1417.333 Million cell updates/sec

Title: US-09-904-987-3

Perfect score: 1435

Sequence: 1 MANLCWMLVLFVATWSDLG.....LFSSPPVILLISFLIFLIVG 253

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 2942292 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*

- 1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
- 2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
- 3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
- 4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
- 5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep:*
- 6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	1435	100.0	253	1 US-08-242-188-2	Sequence 2, Appli
2	1435	100.0	253	1 US-08-509-261A-2	Sequence 2, Appli
3	1435	100.0	253	1 US-08-660-626-8	Sequence 8, Appli
4	1435	100.0	253	1 US-08-692-892-2	Sequence 2, Appli
5	1435	100.0	253	2 US-08-713-939A-2	Sequence 2, Appli
6	1435	100.0	253	2 US-08-868-162A-22	Sequence 22, Appli
7	1435	100.0	253	4 US-09-031-168-8	Sequence 8, Appli
8	1435	100.0	253	4 US-09-128-450-20	Sequence 20, Appli
9	1435	100.0	253	4 US-09-036-579-2	Sequence 2, Appli
10	1435	100.0	253	4 US-09-823-494-20	Sequence 20, Appli
11	1435	100.0	253	4 US-09-550-374-2	Sequence 2, Appli
12	1331.5	92.8	264	4 US-09-128-450-21	Sequence 21, Appli
13	1331.5	92.8	264	4 US-09-823-494-21	Sequence 21, Appli
14	1329	92.6	263	1 US-08-242-188-3	Sequence 3, Appli
15	1329	92.6	263	1 US-08-509-261A-3	Sequence 3, Appli
16	1329	92.6	263	1 US-08-660-626-9	Sequence 9, Appli
17	1329	92.6	263	1 US-08-692-892-3	Sequence 3, Appli
18	1329	92.6	263	2 US-08-713-939A-3	Sequence 3, Appli
19	1329	92.6	263	2 US-08-868-162A-23	Sequence 23, Appli
20	1329	92.6	263	4 US-09-031-168-9	Sequence 9, Appli
21	1329	92.6	263	4 US-09-036-579-3	Sequence 3, Appli
22	1329	92.6	263	4 US-09-550-374-3	Sequence 3, Appli
23	1321	92.1	255	1 US-08-242-188-4	Sequence 4, Appli
24	1321	92.1	255	1 US-08-509-261A-4	Sequence 4, Appli
25	1321	92.1	255	1 US-08-660-626-10	Sequence 10, Appli
26	1321	92.1	255	1 US-08-692-892-4	Sequence 4, Appli
27	1321	92.1	255	2 US-08-713-939A-4	Sequence 4, Appli

28	1321	92.1	255	2 US-08-868-162A-24	Sequence 24, Appli
29	1321	92.1	255	4 US-09-031-168-10	Sequence 10, Appli
30	1321	92.1	255	4 US-09-036-579-4	Sequence 4, Appli
31	1321	92.1	255	4 US-09-550-374-4	Sequence 4, Appli
32	1314.5	91.6	256	4 US-09-128-450-22	Sequence 22, Appli
33	1314.5	91.6	256	4 US-09-823-494-22	Sequence 22, Appli
34	1308.5	91.2	254	4 US-09-128-450-26	Sequence 26, Appli
35	1308.5	91.2	254	4 US-09-823-494-26	Sequence 26, Appli
36	1296.5	90.3	254	1 US-08-242-188-1	Sequence 1, Appli
37	1296.5	90.3	254	1 US-08-509-261A-1	Sequence 1, Appli
38	1296.5	90.3	254	1 US-08-660-626-7	Sequence 7, Appli
39	1296.5	90.3	254	1 US-08-692-892-1	Sequence 1, Appli
40	1296.5	90.3	254	2 US-08-713-939A-1	Sequence 1, Appli
41	1296.5	90.3	254	2 US-08-868-162A-21	Sequence 21, Appli
42	1296.5	90.3	254	4 US-09-031-168-7	Sequence 7, Appli
43	1296.5	90.3	254	4 US-09-128-450-19	Sequence 19, Appli
44	1296.5	90.3	254	4 US-09-036-579-1	Sequence 1, Appli
45	1296.5	90.3	254	4 US-09-823-494-19	Sequence 19, Appli

ALIGNMENTS

RESULT 1

US-08-242-188-2
; Sequence 2, Application US/08242188
; Patent No. 5565186
; GENERAL INFORMATION:
; APPLICANT: Prusiner, Stanley B.
; APPLICANT: Scott, Michael R.
; APPLICANT: Telling, Glenn
; TITLE OF INVENTION: METHOD OF DETECTING PRIONS IN A SAMPLE
; TITLE OF INVENTION: AND TRANSGENIC ANIMAL USED FOR SAME
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Karl Bosicevic
; STREET: 2200 Sand Hill Road
; CITY: Menlo Park
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/242,188
; FILING DATE: 13-MAY-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Bosicevic, Karl
; REGISTRATION NUMBER: 28,807
; REFERENCE/DOCKET NUMBER: 06510/014001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 854-5277
; TELEFAX: (415) 854-0875
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: HUMAN PRION PROTEIN, HuPrp
US-08-242-188-2

Query Match 100.0%; Score 1435; DB 1; Length 253;
Best Local Similarity 100.0%; Pred. No. 2e-133;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MANLCWMLVLFVATWSDLGLCKRKPFGGNTGSRYPGQSPGNGNRYPPQGGGNGQP 60

Db 61 HGGGWCQPHGGGQPHGGGQGGTSHQWKNKPKTKNMHAGAAAAGA 120
QY 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNONNFVHDCV 180
Db 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNONNFVHDCV 180
QY 181 NITIKQHTVTTTNGENFTETDVKMERVVEQMCITQYERESQAYQRGSSMVLSSPPV 240
Db 181 NITIKQHTVTTTNGENFTETDVKMERVVEQMCITQYERESQAYQRGSSMVLSSPPV 240
QY 241 ILLISFLIFLIVG 253
Db 241 ILLISFLIFLIVG 253

RESULT 4

US-08-692-892-2
; Sequence 2, Application US/08692892
; Patent No. 5792901
; GENERAL INFORMATION:
; APPLICANT: Prusiner, Stanley B.
; APPLICANT: Scott, Michael R.
; APPLICANT: Telling, Glenn
; TITLE OF INVENTION: DETECTING PRIONS IN A SAMPLE AND
; TITLE OF INVENTION: PRION PREPARATION AND TRANSGENIC ANIMAL USED FOR SAME
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Karl Bozicevic
; STREET: 2200 Sand Hill Road
; CITY: Menlo Park
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/692,892
; FILING DATE: 30-JULY-1996
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bozicevic, Karl
; REGISTRATION NUMBER: 28,807
; REFERENCE/DOCKET NUMBER: 06510/056001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 322-5070
; TELEFAX: (415) 854-0875
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORGANISM: HUMAN PRION PROTEIN, HuPrP
US-08-692-892-2

Query Match 100.0%; Score 1435; DB 1; Length 253;
Best Local Similarity 100.0%; Pred. No. 2e-133;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MANLGCWMLVLFVATWSDGLCKKRPKPGWNTGGSRYPGQSPGNNRYPPQGGGGWGQ 60
Db 1 MANLGCWMLVLFVATWSDGLCKKRPKPGWNTGGSRYPGQSPGNNRYPPQGGGGWGQ 60
QY 61 HGGGWCQPHGGGQPHGGGQGGTSHQWKNKPKTKNMHAGAAAAGA 120
Db 61 HGGGWCQPHGGGQPHGGGQGGTSHQWKNKPKTKNMHAGAAAAGA 120

QY 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNONNFVHDCV 180
Db 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNONNFVHDCV 180
QY 181 NITIKQHTVTTTNGENFTETDVKMERVVEQMCITQYERESQAYQRGSSMVLSSPPV 240
Db 181 NITIKQHTVTTTNGENFTETDVKMERVVEQMCITQYERESQAYQRGSSMVLSSPPV 240
QY 241 ILLISFLIFLIVG 253
Db 241 ILLISFLIFLIVG 253

RESULT 5

US-08-713-939A-2
; Sequence 2, Application US/08713939A
; Patent No. 5846533
; GENERAL INFORMATION:
; APPLICANT: Prusiner, Stanley B.
; APPLICANT: Williamson, R. Anthony
; APPLICANT: Burton, Dennis R.
; TITLE OF INVENTION: ANTIBODIES SPECIFIC FOR NATIVE PrP
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 2200 Sand Hill Road
; CITY: Menlo Park
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/713,939A
; FILING DATE: 13-SEP-1996
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bozicevic, Karl
; REGISTRATION NUMBER: 28,807
; REFERENCE/DOCKET NUMBER: 06510/059001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-854-5277
; TELEFAX: 415-854-0875
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-713-939A-2

Query Match 100.0%; Score 1435; DB 2; Length 253;
Best Local Similarity 100.0%; Pred. No. 2e-133;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MANLGCWMLVLFVATWSDGLCKKRPKPGWNTGGSRYPGQSPGNNRYPPQGGGGWGQ 60
Db 1 MANLGCWMLVLFVATWSDGLCKKRPKPGWNTGGSRYPGQSPGNNRYPPQGGGGWGQ 60
QY 61 HGGGWCQPHGGGQPHGGGQGGTSHQWKNKPKTKNMHAGAAAAGA 120
Db 61 HGGGWCQPHGGGQPHGGGQGGTSHQWKNKPKTKNMHAGAAAAGA 120
QY 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNONNFVHDCV 180
Db 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNONNFVHDCV 180

QY 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHMRYPNQVYRPMDEYSNQNNFVHDCV 180
|||||
Db 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHMRYPNQVYRPMDEYSNQNNFVHDCV 180
QY 181 NITIKQHTVTTTNGENFTETDVKMERVVEQMCITQYERESQAYYQRGSSMVLFSPPV 240
|||||
Db 181 NITIKQHTVTTTNGENFTETDVKMERVVEQMCITQYERESQAYYQRGSSMVLFSPPV 240
QY 241 ILLISFLIFLIVG 253
|||||
Db 241 ILLISFLIFLIVG 253

RESULT 8

US-09-128-450-20
; Sequence 20, Application US/09128450
; Patent No. 6211149
; GENERAL INFORMATION:
; APPLICANT: Chesebro, Bruce W
; APPLICANT: Caughey, Byron W
; APPLICANT: Chabry, Joelle
; APPLICANT: Priola, Susette
; TITLE OF INVENTION: Inhibitors of Formation of Protease Resistant Prion
; FILE REFERENCE: 50121
; CURRENT APPLICATION NUMBER: US/09/128,450
; CURRENT FILING DATE: 1998-08-03
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 20
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-128-450-20

Query Match 100.0%; Score 1435; DB 4; Length 253;
Best Local Similarity 100.0%; Pred. No. 2e-133;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGSGRYPGQSGPGGNRYPPQGGGGWGQP 60
|||||
Db 1 MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGSGRYPGQSGPGGNRYPPQGGGGWGQP 60
QY 61 HGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGHGTHSQWNKPSKPKTKNMKHMAGAAAAGA 120
|||||
Db 61 HGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGHGTHSQWNKPSKPKTKNMKHMAGAAAAGA 120
QY 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHMRYPNQVYRPMDEYSNQNNFVHDCV 180
|||||
Db 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHMRYPNQVYRPMDEYSNQNNFVHDCV 180
QY 181 NITIKQHTVTTTNGENFTETDVKMERVVEQMCITQYERESQAYYQRGSSMVLFSPPV 240
|||||
Db 181 NITIKQHTVTTTNGENFTETDVKMERVVEQMCITQYERESQAYYQRGSSMVLFSPPV 240
QY 241 ILLISFLIFLIVG 253
|||||
Db 241 ILLISFLIFLIVG 253

RESULT 9

US-09-036-579-2
; Sequence 2, Application US/09036579
; Patent No. 6250954
; GENERAL INFORMATION:
; APPLICANT: Prusiner, Stanley B.
; APPLICANT: Williamson, R. Anthony
; APPLICANT: Burton, Dennis R.
; TITLE OF INVENTION: ANTIBODIES SPECIFIC FOR NATIVE PrP
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.

; STREET: 2200 Sand Hill Road
; CITY: Menlo Park
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/036,579
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/713,939
; FILING DATE: 13-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Bozicevic, Karl
; REGISTRATION NUMBER: 28,807
; REFERENCE/DOCKET NUMBER: 06510/059001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-854-5277
; TELEFAX: 415-854-0875
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-036-579-2

Query Match 100.0%; Score 1435; DB 4; Length 253;
Best Local Similarity 100.0%; Pred. No. 2e-133;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGSGRYPGQSGPGGNRYPPQGGGGWGQP 60
|||||
Db 1 MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGSGRYPGQSGPGGNRYPPQGGGGWGQP 60
QY 61 HGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGHGTHSQWNKPSKPKTKNMKHMAGAAAAGA 120
|||||
Db 61 HGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGHGTHSQWNKPSKPKTKNMKHMAGAAAAGA 120
QY 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHMRYPNQVYRPMDEYSNQNNFVHDCV 180
|||||
Db 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHMRYPNQVYRPMDEYSNQNNFVHDCV 180
QY 181 NITIKQHTVTTTNGENFTETDVKMERVVEQMCITQYERESQAYYQRGSSMVLFSPPV 240
|||||
Db 181 NITIKQHTVTTTNGENFTETDVKMERVVEQMCITQYERESQAYYQRGSSMVLFSPPV 240
QY 241 ILLISFLIFLIVG 253
|||||
Db 241 ILLISFLIFLIVG 253

RESULT 10

US-09-823-494-20
; Sequence 20, Application US/09823494
; Patent No. 6355610
; GENERAL INFORMATION:
; APPLICANT: Chesebro, Bruce W
; APPLICANT: Caughey, Byron W
; APPLICANT: Chabry, Joelle
; APPLICANT: Priola, Susette
; TITLE OF INVENTION: Inhibitors of Formation of Protease Resistant Prion
; FILE REFERENCE: 50121
; CURRENT APPLICATION NUMBER: US/09/823,494
; CURRENT FILING DATE: 2001-03-30


```
;
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/509,261A
; FILING DATE: 31-JUL-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bozicevic, Karl
; REGISTRATION NUMBER: 28,807
; REFERENCE/DOCKET NUMBER: 6510-030001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-327-3400
; TELEFAX: 650 327-3231
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 263 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-08-509-261A-3

Query Match          92.6%; Score 1329; DB 1; Length 263;
Best Local Similarity 87.7%; Pred. No. 5.5e-123;
Matches 228; Conservative 19; Mismatches 5; Indels 8; Gaps 1;

QY  2  ANLGCWMLVLFVATWSDGLCKKRPKPGGNTGSGRYPGGSPGGNRYPPQGGGGWQPH 61
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY  4  SHIGSWILVLFVAMWSDVGLCKKRPKPGGNTGSGRYPGGSPGGNRYPPQGGGGWQPH 63
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY  62 GGGWGQPHGGGGWQPHGGGGWQPHGGGGWQPHGGGGWQPHGGGGWQPHGGGGWQPH 113
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY  64 GGGWGQPHGGGGWQPHGGGGWQPHGGGGWQPHGGGGWQPHGGGGWQPHGGGGWQPH 123
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY  114 GAAAGAVGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGL 173
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY  124 GAAAGAVGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGL 183
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY  174 NFVHDCVNITIKOHTVTTTGTGNETETDVKMMERVVEQMCITQYERESQAYYQGS 233
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY  184 NFVHDCVNITVKEHTVTTTGTGNETETDVKMMERVVEQMCITQYERESQAYYQGS 243
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY  234 LFSSPPVILLISFLIFLVG 253
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:
QY  244 LFSSPPVILLISFLIFLVG 263
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:

```

Search completed: January 28, 2003, 14:21:21
Job time : 6.25212 secs

Result No.	Score	Query			DB	ID	Description
		Match	Length	Count			
1	702	100.0	140	2	A49669	synuclein alpha -	
2	485	69.1	149	2	B43959	synuclein, form 2	
3	409	58.3	134	2	S44330	synuclein - human	
4	408	58.1	134	2	S39046	phosphonuroprotei	
5	407	58.0	137	2	S15648	phosphonuroprotei	
6	273.5	39.0	123	2	A57431	synuclein-like pro	
7	261	37.2	113	2	A60887	synuclein - pacifi	
8	245.5	35.0	124	2	A60218	14K brain-specific	
9	129	18.4	28	2	S66746	alpha-synuclein, N	
10	123	17.5	150	2	T37479	protein 135, daue	
11	122	17.4	798	2	T21369	hypothetical prote	
12	105	15.0	643	1	T07064	seed biotin-contai	
13	100	14.2	317	2	S33616	embryonic abundant	
14	94	13.4	224	2	S16259	embryonic abundant	
15	90.5	12.9	1612	2	A81347	probable peptidogl	
16	89	12.7	474	2	T01347	pectinesterase hom	
17	89	12.7	739	2	H75001	methyl-accepting c	
18	89	12.7	1109	2	A56143	surface-array prot	
19	88.5	12.6	322	2	A57030	membrane protease	
20	88	12.5	464	2	A32461	CAMP-dependent pro	
21	87.5	12.5	473	2	S61428	embryonic abundant	
22	87.5	12.5	474	2	S39475	embryonic protein	
23	87.5	12.5	742	1	SYBS2G	phosphoribosylform	
24	87	12.4	424	2	B70365	phosphoribosylamin	
25	87	12.4	469	2	E81121	phage sheath prote	
26	87	12.4	817	2	JG4176	pyruvate, water di	
27	87	12.4	1093	2	AC1753	tail protein [bact	
28	86	12.3	739	2	F71161	probable chemotaxi	
29	86	12.3	1785	2	A45546	major merozoite su	

[illegible]

Db 66 AAEKTESGIAFVGQKIAEGAGAVAGG-----AKAAGGAVVDGI 103

RESULT 11

T071369
hypothetical protein F25H8.5a - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T21369
R:Gajadaty, S.
submitted to the EMBL Data Library, February 1996
A:Reference number: Z19413
A:Accession: T21369
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-796 <WIL>
A:Cross-references: EMBL:Z69360; PIDN:CAA93285.1; GSPDB:GN00022; CESP:F25H8.5a
A:Experimental source: clone F25H8
C:Genetics:
A:Gene: CESP:F25H8.5a
A:Map position: 4
A:Introns: 23/3; 43/1; 303/1; 321/1; 343/1; 380/1; 387/1; 405/1; 427/1; 449/1; 478/1; 49

Query Match 17.4%; Score 122; DB 2; Length 798;
Best Local Similarity 32.7%; Pred. No. 0.035;
Matches 35; Conservative 13; Mismatches 39; Indels 20; Gaps 4;

Qy 13 EGVVAAAEKTKQGVAAEAGKTKGVLVYVSGTKGVLVGVVATVAETK---EQVTNVGGA 69

Db 610 DGAVAAKDAAEVAGATTAGNKTVEGMRFG-----ENVAHGAGKVADGAKAAGDTVVEGGA 665

Qy 70 VV---TGVTAVAKTVEGAGSTAAATGFVKDQGLKNEEAPQEG 112

Db 666 AAEKTESGIAFVGQKIAEGAGAVAGG-----AKAAGGAVVDGV 703

RESULT 12

T07064
seed biotin-containing protein LEA [validated] - soybean
C:Species: Glycine max (soybean)
C:Date: 01-Sep-2000 #sequence_revision 01-Sep-2000 #text_change 01-Feb-2002
C:Accession: T07064
R:Hsing, Y.C.; Tsou, C.H.; Hsu, T.F.; Chen, Z.Y.; Hsieh, K.L.; Hsieh, J.S.; Chow, T.Y.
Plant Mol. Biol. 38, 461-490, 1998
A:Title: Tissue- and stage-specific expression of a soybean (Glycine max L.) seed-mature
A:Reference number: Z15895; MUID:98418627; PMID:9747855
A:Accession: T07064
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-643 <HSI>
A:Cross-references: EMBL:U59626; NID:gl389896; PIDN:AAC61783.1; PID:gl389897
A:Experimental source: strain Shi-Shi; cotyledon
C:Superfamily: pea seed biotin-containing protein
C:Keywords: biotin binding; seed
F:125/Binding site: biotin (Lys) (covalent) #status predicted

Query Match 15.0%; Score 105; DB 1; Length 643;
Best Local Similarity 28.2%; Pred. No. 0.65;
Matches 48; Conservative 10; Mismatches 68; Indels 44; Gaps 7;

Qy 2 DVFMKGLSKAKEGVVAAAEKTKQGVAAEAGKTKG-----VLYVGSK---TKEGVVG--- 51

Db 289 DTISSAAKTAASEKTAPVAEAKADYTLQAEKAKSAGGTTTASYGKAVQAKDVAVESGKS 348

Qy 52 ----VATVAEKTKEQVTNVGGA-----VVTGVTAVAKTVEGAG----- 86

Db 349 AAGYAAKVAADLRDKATAVGAAAHFAEKTVEGTAKAAH-VVEGAGYAGHKAELASM 407

Qy 87 -----STAAATGFVKDQGLKNEEAPQEGILEMDPVDPNFAYEMPSE 130

Db 406 SAGAVKGLAASAGETAYETAKKKEAQRE--LEAKKPSQPQAEERPSE 455

RESULT 13

S33616
embryonic abundant protein, group 3 - wheat (fragment)
N:Alternate names: group 3 late embryogenesis abundant protein
C:Species: Triticum aestivum (common wheat)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 29-Oct-1999
C:Accession: S33616; S27767
R:Curry, J.; Walker-Simmons, M.K.
Plant Mol. Biol. 21, 907-912, 1993
A:Title: Unusual sequence of group 3 LEA (II) mRNA inducible by dehydration stress in
A:Reference number: S33616; MUID:93222487; PMID:8467082
A:Accession: S33616
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-317 <WAL>
A:Cross-references: EMBL:M72395; NID:gl70691; PIDN:AAA34267.1; PID:gl70692

Query Match 14.2%; Score 100; DB 2; Length 317;
Best Local Similarity 29.6%; Pred. No. 0.75;
Matches 53; Conservative 10; Mismatches 62; Indels 54; Gaps 8;

Qy 6 KCLSKAKEGVVAAAEKTKQGVAAE-----AAGKTKGVLVYVSGTKGVLVGVVATVAE 57

Db 9 EKEGASEMADAAGKTKDAAAAEKTRMADAAAGKTKETKDAAVEKTRMADTAATKAA 67

Qy 58 KTK-----BQVT-----NVGGAVT-----GVTAQAQTKVEGA--- 85

Db 68 ETKDAAEKASGAGEMVTETKASAKDAADKASGAETVTEKAKGAKDAALDTAEGAKEY 127

Qy 86 -----GSTAATGVKKDQGLKNEEAPQEGILEMDPVDPNFAYEMPSEEGYQDEP 138

Db 128 MYDKKEDARRALAGSV-KDSKGETNESACQGG-----QDVRRAAEKAEARQTHOP 179

RESULT 14

S16259
embryonic abundant protein, group 3 - common wheat x Sanduri wheat
N:Alternate names: late embryogenesis abundant protein
C:Species: Triticum aestivum x Triticum timopheevi (common wheat x Sanduri wheat)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 02-Sep-2000
C:Accession: S16259
R:Curry, J.; Morris, C.F.; Walker-Simmons, M.K.
Plant Mol. Biol. 16, 1073-1076, 1991
A:Title: Sequence analysis of a cDNA encoding a Group 3 LEA mRNA inducible by ABA or
A:Reference number: S16259; MUID:91322498; PMID:1830822
A:Accession: S16259
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-224 <CUR>
A:Cross-references: EMBL:X58882; NID:921636; PIDN:CAA40204.1; PID:g21637
C:Superfamily: barley hypothetical protein (clone ES2A)

Query Match 13.4%; Score 94; DB 2; Length 224;
Best Local Similarity 30.2%; Pred. No. 1.6;
Matches 35; Conservative 12; Mismatches 55; Indels 14; Gaps 4;

Qy 6 KGLSKAKEGVVAAAEKTKQG-----VAEAGKTKGVLVYVSGTKGVLVGVVATVAETK 60

Db 16 KARNEEKTGVGMATKDKAGQTTATKQAGETTEATKQAAETTEAAKQKASETAEATK 75

Qy 61 EQVTNVGAVTGVTAQAQKTEVAGAGSI--AAATGFVKDQ-----LGKNEGAPOE 110

Db 76 QKAAEAKDKTAQTAQAQKTEVAGAGSI--AAATGFVKDQ-----KDQTASTLGKTEAAKQK 128

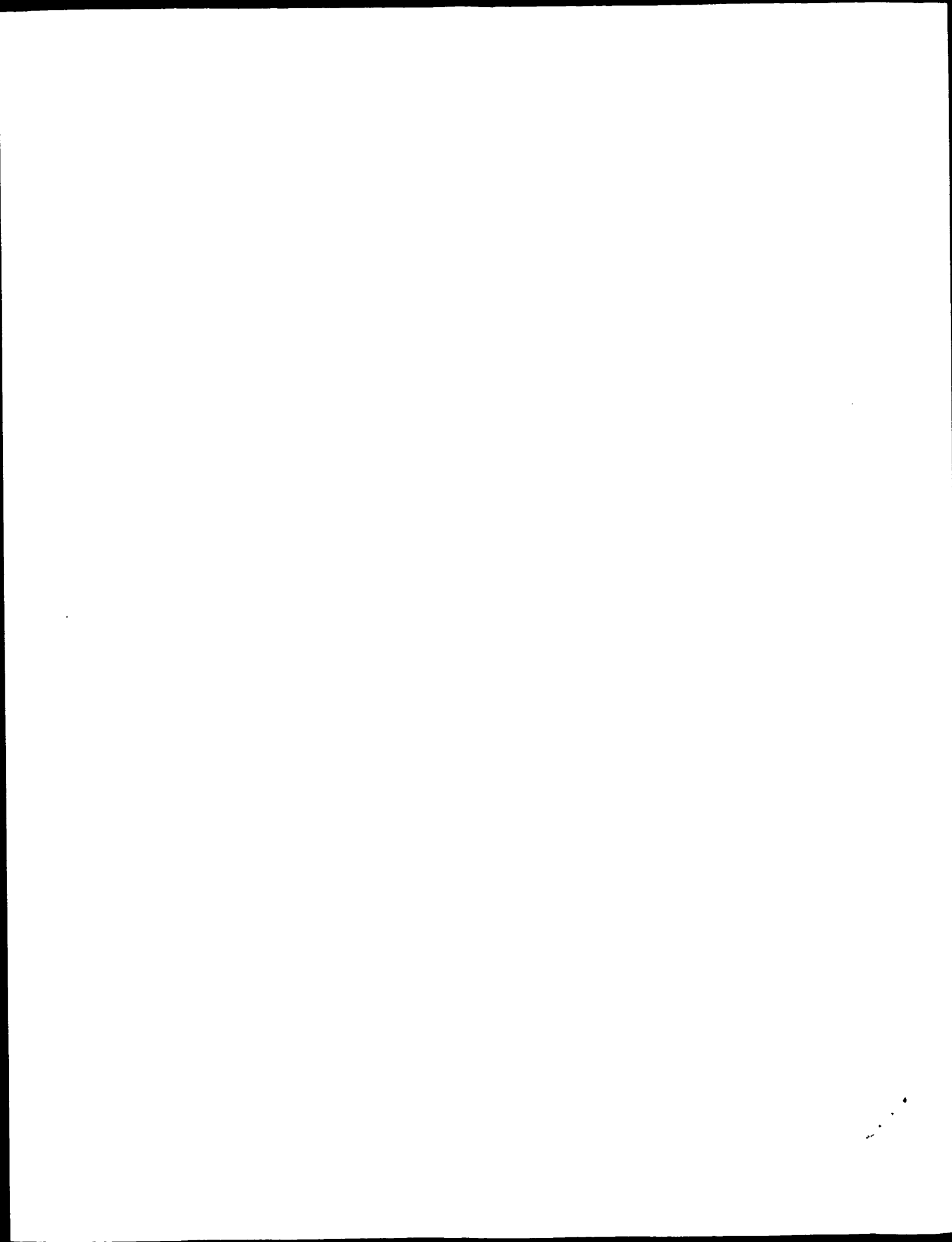
RESULT 15

AB1347
probable peptidoglycan bound protein (LPXTG motif) lmo2178 [imported] - Listeria mono
C:Species: Listeria monocytogenes
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C:Accession: AB1347
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi,

D.: Jones, L.M.: Karst, U.
Science 294, 849-852, 2001
A:Authors: Kref, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wenland,
A:Title: Comparative genomics of *Listeria* species
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AB1347
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-1612 <GLA>
A:Cross-references: GB:NC_003210; PIDN:CAD00256.1; PID:gl6411648; GSPDB:GN00177
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: lmo2178

Query Match 12.9%; Score 90.5; DB 2; Length 1612;
Best Local Similarity 28.1%; Pred. No. 26;
Matches 45; Conservative 17; Mismatches 65; Indels 33; Gaps 7;
QY 3 VFMKGLSKAKGVAA-----AEKTKQG---VAEAAAGK-----TKEGVL 38
Db 1110 VLLKEDSATKDAIAGAEFELQNAAGTKVADNLVSNADGKIEVTDLAPGDYQFVETKAPTG 1169
QY 39 YV--GSKTKEGVYHGVATVAEKTKEQVTNVGGAVVTGVTVAQKTVGAG-SIAAATGFV 95
Db 1170 YVLDGAPTKEVFEFNOEAAVITKENTAKTGSVLTKEDSVSKATISGAEEFELQNAAGTK 1229
QY 96 KKDQLGKNEGAPQEGILEDMPVDP-DNEAYEMPSEEGYQ 134
Db 1230 VKDNLTTNADGK-----LEVTDLAPGDYKFVETKAPTGYE 1264

Search completed: January 28, 2003, 14:20:02
Job time : 5.67113 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:03:18 ; Search time 2.14149 Seconds
(without alignments)
2711.515 Million cell updates/sec

Title: US-09-904-987-4

Perfect score: 702

Sequence: 1 MDVFMKGLSKAGEYVAAAE.....DNEAYEMPSEGYQDIPEEA 140

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	702	100.0	140	1 SYUA_HUMAN	P37840 homo sapien
2	670	95.4	140	1 SYUA_RAT	P37377 rattus norv
3	669	95.3	140	1 SYUA_MOUSE	O55042 mus musculu
4	595.5	84.8	143	1 SYUA_POEGU	Q91448 poephila gu
5	409	58.3	134	1 SYUB_HUMAN	Q16143 homo sapien
6	408	58.1	134	1 SYUB_BOVIN	P33567 bos taurus
7	407	58.0	137	1 SYUB_RAT	Q63754 rattus norv
8	377	48.0	143	1 SYU_TORCA	P37379 torpedo cal
9	313	44.6	127	1 SYUG_HUMAN	O76070 homo sapien
10	312.5	44.5	123	1 SYUG_MOUSE	Q920f7 mus musculu
11	307.5	43.8	127	1 SYUG_HUMAN	Q9nz50 homo sapien
12	273.5	39.0	123	1 SYUG_RAT	O63544 rattus norv
13	94	13.4	224	1 LEA3_WHEAT	Q03968 triticum ae
14	88	12.5	464	1 AK15_RAT	Q03968 triticum ae
15	87.5	12.5	742	1 PURL_BACSU	P24587 rattus norv
16	87	12.4	424	1 PURL_AQUAE	P12042 bacillus su
17	87	12.4	817	1 PPSA_PYREF	O66949 aquifex ae
18	86.5	12.3	524	1 P60_LISIV	Q42850 pyrococcus
19	85	12.1	791	1 K6PP_RABIT	Q01837 listeria iv
20	85	12.1	2038	1 FSH_DROME	P47859 ocyctolaqu
21	84.5	12.0	524	1 P60_LISWE	P13709 drosophila
22	82.5	11.8	280	1 LE76_BRANA	Q01839 listeria we
23	82.5	11.8	1528	1 SPAA_STRDO	P13934 brassica na
24	82	11.7	321	1 EAEB_ECO27	P1979 streptococc
25	82	11.7	433	1 HPR2_HALVA	Q05129 escherichia
26	82	11.7	3099	1 POLG_PEMVM	P42258 haloarcula
27	81.5	11.6	397	1 LE11_THETN	O56075 p genome po
28	81	11.5	682	1 AMPH_CHICK	Q8rdk3 thermocanaer
29	80.5	11.5	257	1 NT6A_HUMAN	P50478 gallus gall
30	79.5	11.3	129	1 YD39_HAEIN	P34132 homo sapien
31	79.5	11.3	819	1 PPSA_PYRAB	P71378 haemophilus
32	79	11.3	201	1 DRPF_CRAPL	Q9v2h7 pyrococcus
33	79	11.3	570	1 SYE_PYRHO	P32883 craterostig

RESULT 1

ID	SYUA_HUMAN	STANDARD;	PRT;	140 AA.
AC	P37840; Q13701;			
DT	01-OCT-1994 (Rel. 30, Created)			
DT	01-OCT-1994 (Rel. 30, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Alpha-synuclein (Non-A beta component of AD amyloid) (Non-A4 component of amyloid precursor) (NACP).			
DE	SNCA OR NACP.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
OX	[1]			
RP	SEQUENCE FROM N.A. (ISOFORM 1), AND SEQUENCE OF 61-95.			
RC	TISSUE=Brain;			
RX	MEDLINE=94068588; PubMed=8248242;			
RA	Ueda K., Fukushima H., Masliak E., Xia Y., Iwai A., Yoshimoto M., Otero D.A., Kondo J., Ihara Y., Saitoh T.;			
RA	"Molecular cloning of cDNA encoding an unrecognized component of amyloid in Alzheimer disease.";			
RT	Proc. Natl. Acad. Sci. U.S.A. 90:11282-11286(1993).			
RL	[2]			
RN	SEQUENCE FROM N.A. (ISOFORMS 1; 2-4 AND 2-5).			
RP	MEDLINE=95324916; PubMed=7601450;			
RX	Campion D., Martin C., Heilig R., Charbonnier F., Moreau V., Flaman J.-M., Petit J.-L., Hannequin D., Brice A., Frebourg T.;			
RA	"The NACP/synuclein gene: chromosomal assignment and screening for alterations in Alzheimer disease.";			
RT	Genomics 26:254-257(1995).			
RL	[3]			
RN	SEQUENCE FROM N.A. (ISOFORM 2-4).			
RP	TISSUE=Brain;			
RC	MEDLINE=95100969; PubMed=7802671;			
RX	Ueda K., Saitoh T., Mori H.;			
RA	"Tissue-dependent alternative splicing of mRNA for NACP, the precursor of non-A beta component of Alzheimer's disease amyloid.";			
RT	Biochem. Biophys. Res. Commun. 205:1366-1372(1994).			
RL	[4]			
RN	SEQUENCE FROM N.A.			
RP	Xia Y., Silva R.D., Chen X.H., Saitoh T.;			
RA	Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.			
RL	[5]			
RN	PHOSPHORYLATION BY CASEIN KINASE.			
RP	MEDLINE=20085057; PubMed=10617630;			
RX	Okochi M., Walter J., Koyama A., Nakajo S., Baba M., Iwatsubo T., Meijer L., Kahle P.J., Haass C.;			
RA	"Constitutive phosphorylation of the Parkinson's disease associated alpha-synuclein.";			
RT	J. Biol. Chem. 275:390-397(2000).			
RL	[6]			
RN	PHOSPHORYLATION BY G-PROTEIN COUPLED RECEPTOR KINASE.			
RP	MEDLINE=20409007; PubMed=10852916;			
RX	Pronin A.N., Morris A.J., Surguchov A., Benovic J.L.;			
RA	"Synucleins are a novel class of substrates for G protein-coupled			

34	79	11.3	625	1	ILVB_MYCLE	O33112 mycobacteri
35	79	11.3	964	1	LON2_MAIZE	P36486 zea mays (m
36	78.5	11.2	739	1	PURL_LISIN	Q92an9 listeria in
37	78.5	11.2	739	1	PURL_LISMO	Q9y6c1 listeria in
38	77.5	11.0	213	1	LEA1_HORVU	P14928 hordeum vul
39	77.5	11.0	216	1	MBCL_METMA	P58979 methanosarc
40	77.5	11.0	221	1	LEA3_MAIZE	O42376 zea mays (m
41	77.5	11.0	280	1	MACS_CHICK	P16527 gallus gall
42	77.5	11.0	503	1	NRFL_HUMAN	Q16656 homo sapien
43	77.5	11.0	503	1	NRFL_MOUSE	Q9wu00 mus musculu
44	77.5	11.0	729	1	PURL_STAM	Q99v28 staphylococ
45	77	11.0	306	1	YOPD_YEREN	P37132 yersinia en

ALIGNMENTS


```

CC -1: PTM: PHOSPHORYLATED, PREDOMINANTLY ON SERINE RESIDUES.
CC PHOSPHORYLATION BY CK1 APPEARS TO OCCUR ON RESIDUES DISTINCT FROM
CC THE RESIDUE PHOSPHORYLATED BY OTHER KINASES.
CC -1: SIMILARITY: BELONGS TO THE SYNUCLEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; S73007; AAB20688.1; -
CC DR EMBL; S73008; AAB20689.1; -
CC DR EMBL; S73009; AAB20690.1; -
CC DR EMBL; AF007758; AAC16026.1; -
CC DR PIR; A43959; A43959.
CC DR PIR; B43959; B43959.
CC DR PIR; C43959; C43959.
CC DR InterPro; IPR001058; Synuclein.
CC DR PRINTS; PR01211; SYNUCLEIN.
CC DR ProDom; PD010631; synuclein; 1.
CC KW Alternative splicing; Repeat; Phosphorylation.
CC FT DOMAIN 20 67
CC FT [EQ]-[GQ]-V-X(4).
CC FT REPEAT 20 30 1.
CC FT REPEAT 31 41 2.
CC FT REPEAT 42 56 3 (APPROXIMATE).
CC FT REPEAT 57 67 4.
CC FT MOD_RES 129 129
CC FT PHOSPHORYLATION (BY CK2, GRK2 AND GRK5)
CC FT (BY SIMILARITY).
CC FT S -> R (IN ISOFORM SYN3).
CC FT MISSING (IN ISOFORM SYN3).
CC FT EGYEQEILEDMPDPSSEAYEMPEEGYQDYEPFA ->
CC FT YPMGECTNHPRLIALRVKRSYREHSWRPRKQLSLACVMD
CC FT PFLPT (IN ISOFORM SYN2).
CC FT SEQUENCE 140 AA; 14515 MW; 1FFD19CD3B9E636C CRC64;
CC SQ
CC
CC Query Match 95.4%; Score 670; DB 1; Length 140;
CC Best Local Similarity 95.0%; Pred. No. 1e-44;
CC Matches 133; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
CC
CC QY 1 MDVFMKGLSKAKGVVAAAEKTKQGVAAAGTKGVLVVGSKTKGVVGHVATVAEKT 60
CC |
CC Db 1 MDVFMKGLSKAKGVVAAAEKTKQGVAAAGTKGVLVVGSKTKGVVGHVATVAEKT 60
CC |
CC QY 61 EQVTNCGAVVTGTTAAVAKTVEGAGSIAAATGFVKKQDLGKNEGAPQEGILEDMPDVP 120
CC |
CC Db 61 EQVTNCGAVVTGTTAAVAKTVEGAGSIAAATGFVKKQDLGKNEGAPQEGILEDMPDVP 120
CC |
CC QY 121 DNEAYEMPEEGYQDYEPFA 140
CC |
CC Db 121 SSEAYEMPEEGYQDYEPFA 140
CC |
CC RESULT 3
CC ID SYUA_MOUSE STANDARD; PRT; 140 AA.
CC AC O55042; Q9QUR3;
CC DT 15-JUL-1999 (Rel. 38, Created)
CC DT 16-OCT-2001 (Rel. 40, Last sequence update)
CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
CC DE Alpha-synuclein (Non-A beta component of AD amyloid) (Non-A4 component
CC DE of amyloid precursor) (NACP).
CC GN SNCA OR SYN.
CC OS Mus musculus (Mouse).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
CC OX NCBI_TaxID=10090;
CC RN [1]
CC RP SEQUENCE FROM N.A.

```

```

RX MEDLINE=98311217; PubMed=9648883;
RA Hsu L.J., Mallory M., Xia Y., Weinbergs I., Hashimoto M.,
RA Yoshimoto M., Thal L.J., Saitoh T., Masliah E.;
RT "Expression pattern of synucleins (non-Abeta component of Alzheimer's
RT disease amyloid precursor protein/alpha-synuclein) during murine
RT brain development.";
RL J. Neurochem. 71:338-344(1998).
RN [2]
RP SEQUENCE FROM N.A.
RP STRAIN=ICR;
RX MEDLINE=98264007; PubMed=9601701;
RA Hong L., Ko H.W., Gwag B.J., Joe E., Lee S., Kim Y.T., Suh Y.-H.;
RA "The cDNA cloning and ontogeny of mouse alpha-synuclein.";
RL NeuroReport 9:1239-1243(1998).
RN [3]
RP SEQUENCE FROM N.A.
RP STRAIN=BL6, and SV129/OLA; TISSUE=Brain;
RA Fog J.U., Kallunki P.;
RT "Genomic cloning of the mouse alpha-synuclein and analysis of the
RT promoter.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RP STRAIN=129/SvJ;
RA Touchman J.W., Dehejia A., Chiba-Falek O., Cabin D.E., Schwartz J.R.,
RA Orrison B.M., Polymopoulos M.H., Nussbaum R.D.;
RT "Human and mouse alpha-synuclein genes: comparative genomic sequence
RT analysis and identification of a novel gene regulatory element.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
CC -1: FUNCTION: MAY BE INVOLVED IN NEURONAL PLASTICITY.
CC -1: PTM: PHOSPHORYLATED, PREDOMINANTLY ON SERINE RESIDUES (BY
CC SIMILARITY).
CC -1: SIMILARITY: BELONGS TO THE SYNUCLEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF044672; AAC00521.1; -
CC DR EMBL; AF033261; AAD11254.1; -
CC DR EMBL; AF179273; AAD56908.1; -
CC DR EMBL; AF179272; AAD56907.1; -
CC DR EMBL; AF179268; AAD56907.1; JOINED.
CC DR EMBL; AF179269; AAD56907.1; JOINED.
CC DR EMBL; AF179270; AAD56907.1; JOINED.
CC DR EMBL; AF179271; AAD56907.1; JOINED.
CC DR EMBL; AF163865; AAG30304.1; -
CC DR MGB; MGI:1277151; Snca.
CC DR InterPro; IPR001058; Synuclein.
CC DR Pfam; PF01387; Synuclein; 1.
CC DR PRINTS; PR01211; SYNUCLEIN.
CC DR ProDom; PD010631; Synuclein; 1.
CC KW Repeat; Phosphorylation.
CC FT DOMAIN 20 67
CC FT 4 X 11 AA TANDEM REPEATS OF [EGS]-K-T-K-
CC FT [EQ]-[GQ]-V-X(4).
CC FT REPEAT 20 30 1.
CC FT REPEAT 31 41 2.
CC FT REPEAT 42 56 3 (APPROXIMATE).
CC FT REPEAT 57 67 4.
CC FT MOD_RES 129 129
CC FT PHOSPHORYLATION (BY CK2, GRK2 AND GRK5)
CC FT (BY SIMILARITY).
CC FT K -> T (IN REF. 1).
CC FT CONFLICT 58 58
CC SQ SEQUENCE 140 AA; 14485 MW; 1FFD19D7E15E636C CRC64;
CC
CC Query Match 95.3%; Score 669; DB 1; Length 140;
CC Best Local Similarity 95.0%; Pred. No. 1.2e-44;
CC Matches 133; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
CC
CC QY 1 MDVFMKGLSKAKGVVAAAEKTKQGVAAAGTKGVLVVGSKTKGVVGHVATVAEKT 60

```

```

Db 1 MDVFMKLSKAKGVAAAEKTKQGVAAAGKTKGVLVYVGSKTKEGVHGVTTVAEKT 60
QY 61 EQVTVNGGAVVTVTAQAQKTVGAGSIAAATGFKVKKDQKNGEAGAPQGGILEDMVDP 120
Db 61 EQVTVNGGAVVTVTAQAQKTVGAGSIAAATGFKVKKDQKNGEAGAPQGGILEDMVDP 120
QY 121 DNEAYEMPEEGYQDYEP 140
Db 121 GSEAYEMPEEGYQDYEP 140

RESULT 4
SYUA_POEGU
ID SYUA_POEGU STANDARD; PRT; 143 AA.
AC Q91448;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Alpha-synuclein (Synelfin).
OS Poephila guttata (zebra finch) (Taeniopygia guttata).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Passeriformes; Estrildidae;
OC Estrildinae; Taeniopygia.
OX NCBI_TaxID=59729;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA MEDLINE=95374791; PubMed=7646890;
RT "Characterization of a novel protein regulated during the critical
RT period for song learning in the zebra finch."
RL Neuron 15:361-372(1995).
CC -1- FUNCTION: MAY SERVE A FUNCTION CRITICAL TO THE REGULATION OF
CC VERTEBRATE NEURAL PLASTICITY.
CC -1- TISSUE SPECIFICITY: BRAIN.
CC -1- SIMILARITY: BELONGS TO THE SYNUCLEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L33860; AAA93538.1; -.
DR InterPro: IPR001038; Synuclein.
DR Pfam: PF01387; Synuclein; 1.
DR PRINTS: PR01211; SYNUCLEIN.
DR PRODOM: PD010631; Synuclein; 1.
KW Repeat.
FT DOMAIN 20 67 4 X 11 AA TANDEM REPEATS OF [EGS]-K-T-K-
FT REPEAT 20 30 1.
FT REPEAT 31 41 2.
FT REPEAT 42 56 3 (APPROXIMATE).
FT REPEAT 57 67 4.
SQ SEQUENCE 143 AA; 14874 MW; 9C3AB9C9902C54A7 CRC64;

Query Match 84.8%; Score 595.5; DB 1; Length 143;
Best Local Similarity 86.7%; Pred. No. 4.9e-39;
Matches 124; Conservative 6; Mismatches 10; Indels 3; Gaps 2;

QY 1 MDVFMKLSKAKGVAAAEKTKQGVAAAGKTKGVLVYVGSKTKEGVHGVTTVAEKT 60
Db 1 MDVFMKLSKAKGVAAAEKTKQGVAAAGKTKGVLVYVGSKTKEGVHGVTTVAEKT 60
QY 61 EQVTVNGGAVVTVTAQAQKTVGAGSIAAATGFKVKKDQKNGEAGAPQGGILEDMVDP 117
Db 61 EQVTVNGGAVVTVTAQAQKTVGAGSIAAATGFKVKKDQKNGEAGAPQGGILEDMVDP 120
QY 118 VDPNEAYEMPEEGYQDYEP 140

```

```

Db 121 VDPNEAYEMPEEGYQDYEP 143
RESULT 5
SYUB_HUMAN
ID SYUB_HUMAN STANDARD; PRT; 134 AA.
AC Q16143;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Beta-synuclein.
GN SNCB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=94252398; PubMed=8194594;
RA Jakes R., Spillantini M.G., Goedert M.;
RT "Identification of two distinct synucleins from human brain."
RL FEBS Lett. 345:27-32(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99026142; PubMed=9806846;
RA Lavedan C., Leroy E., Torres R., Dehejia A., Dutra A., Buchholz S.,
RA Nussbaum R.L., Polymeropoulos M.H.;
RT "Genomic organization and expression of the human beta-synuclein gene
RT (SNCB)."
RL Genomics 54:173-175(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP PHOSPHORYLATION.
RX MEDLINE=20409007; PubMed=10852916;
RA Pronin A.N., Morris A.J., Surguchov A., Benovic J.L.;
RT "Synucleins are a novel class of substrates for G protein-coupled
RT receptor kinases."
RL J. Biol. Chem. 275:26515-26522(2000).
CC -1- FUNCTION: MAY BE INVOLVED IN NEURONAL PLASTICITY.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN BRAIN; CONCENTRATED
CC IN PRESYNAPTIC NERVE TERMINALS.
CC -1- PTM: PHOSPHORYLATED. PHOSPHORYLATION BY G-PROTEIN COUPLED RECEPTOR
CC KINASES (GKR) IS MORE EFFICIENT THAN PHOSPHORYLATION BY CK1, CK2
CC AND CAM-KINASE II.
CC -1- DISEASE: BRAIN IRON ACCUMULATION TYPE 1 (NB1AL, ALSO CALLED
CC HALLERVORDEN-SPATZ SYNDROME), A RARE NEUROAXONAL DYSTROPHY, IS
CC HISTOLOGICALLY CHARACTERIZED BY AXONAL SPHEROIDS, IRON DEPOSITION,
CC LEWY BODY (LB)-LIKE INTRANEUROAL INCLUSIONS, GLIAL INCLUSIONS AND
CC NEUROFIBRILLARY TANGLES. SNCB IS FOUND IN SPHEROIDS BUT NOT IN
CC INCLUSIONS.
CC -1- SIMILARITY: BELONGS TO THE SYNUCLEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: S69965; AAB30860.1; -.
DR EMBL: AF053136; AAC80286.1; -.
DR EMBL: AF053134; AAC80286.1; JOINED.
DR EMBL: AF053135; AAC80286.1; JOINED.
DR EMBL: BC02902; AAC02902.1; -.
DR Genew; HGNC:11140; SNCB.

```

```

DR MIM; 602569;
DR InterPro: IPR001058; Synuclein.
DR Pfam: PF01387; Synuclein; 1.
DR PRINTS: PR01211; SYNUCLEIN.
DR ProDom: PD010631; Synuclein; 1.
KW Phosphorylation; Repeat.
FT DOMAIN 20 67
FT REPEAT 20 30
FT REPEAT 31 41
FT REPEAT 42 56
FT REPEAT 57 67
FT MOD_RES 118 118
SQ SEQUENCE 134 AA; 14288 MW; 5BCA9FCA615AC4EF CRC64;

Query Match 58.3%; Score 409; DB 1; Length 134;
Best Local Similarity 61.6%; Pred. No. 7.4e-25;
Matches 90; Conservative 20; Mismatches 18; Indels 18; Gaps 4;

QY 1 MDVFMKGLSKAKGVVAAAEKTKQGVAAAGTKQGVLYGSKTKGVVGVHGVATVAEKT 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 MDVFMKGLSKAKGVVAAAEKTKQGVTEAAEKTQGVLYGSKTKGVVGVHGVATVAEKT 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 EQVTNNGGAVVGTAVAKTVEGAGSIAAATGFVKKQDLG---KNEE---GAPOEGILE 114
   || :||| : ||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 61 EQASHLGGAVFS-----GAGNIAAATGLVKREEFTDLKPEEVAQEAEEPLIE 109
   : :|| :||| :||| :||| :||| :||| :||| :||| :||| :|||

QY 115 DMPVDPDNEAYEMPSEEGYQDYPEA 140
   : :|| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 110 PL-MEPEGESYEDPPQEEYQYEPEA 134
   : :|| :||| :||| :||| :||| :||| :||| :||| :||| :|||

RESULT 6
SYUB_BOVIN STANDARD; PRT; 134 AA.
AC F33567;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Beta-synuclein (Phosphonucleoprotein 14) (PNP 14) (14 kDa brain-
specific protein).
GN SNCB.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=94039126; PubMed=8223629;
RA Nakajo S., Tsukada K., Omata K., Nakamura Y., Nakaya K.;
RT "A new brain-specific 14-kDa protein is a phosphoprotein. Its
complete amino acid sequence and evidence for phosphorylation.";
RL Eur. J. Biochem. 217:1057-1063(1993).
RN [2]
RP PARTIAL SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=91038084; PubMed=2230807;
RA Nakajo S., Omata K., Aiuchi T., Shibayama T., Okahashi I., Ochiai H.,
RA Nakai Y., Nakaya K., Nakamura Y.;
RT "Purification and characterization of a novel brain-specific 14-kDa
protein.";
RL J. Neurochem. 55:2031-2038(1990).
CC -!- FUNCTION: MAY BE INVOLVED IN NEURONAL PLASTICITY.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- TISSUE SPECIFICITY: SPECIFICALLY PRESENT IN SYNAPSES AROUND
NEURONS BUT NOT IN GLIAL CELLS.
CC -!- PTM: PHOSPHORYLATED. PHOSPHORYLATION BY G-PROTEIN COUPLED RECEPTOR
KINASES (GKR) IS MORE EFFICIENT THAN PHOSPHORYLATION BY CK1, CK2
AND CAM-KINASE II (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE SYNUCLEIN FAMILY.
DR PIR; A60218; A60218.

```

```

DR PIR; JU0238; JU0238.
DR PIR; S39046; S39046.
DR InterPro: IPR001058; Synuclein.
DR Pfam: PF01387; Synuclein; 1.
DR PRINTS: PR01211; SYNUCLEIN.
DR ProDom: PD010631; Synuclein; 1.
KW Phosphorylation; Repeat.
FT DOMAIN 20 67
FT REPEAT 20 30
FT REPEAT 31 41
FT REPEAT 42 56
FT REPEAT 57 67
FT MOD_RES 118 118
SQ SEQUENCE 134 AA; 14277 MW; 484FA0IA01979966 CRC64;

Query Match 58.1%; Score 408; DB 1; Length 134;
Best Local Similarity 62.3%; Pred. No. 8.9e-25;
Matches 91; Conservative 18; Mismatches 19; Indels 18; Gaps 4;

QY 1 MDVFMKGLSKAKGVVAAAEKTKQGVAAAGTKQGVLYGSKTKGVVGVHGVATVAEKT 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 MDVFMKGLSKAKGVVAAAEKTKQGVTEAAEKTQGVLYGSKTKGVVGVHGVATVAEKT 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 EQVTNNGGAVVGTAVAKTVEGAGSIAAATGFVKKQDLG---KNEE---GAPOEGILE 114
   || :||| : ||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 61 EQASHLGGAVFS-----GAGNIAAATGLVKREEFTDLKPEEVAQEAEEPLIE 109
   : :|| :||| :||| :||| :||| :||| :||| :||| :||| :|||

QY 115 DMPVDPDNEAYEMPSEEGYQDYPEA 140
   : :|| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 110 PL-MEPEGESYEDPPQEEYQYEPEA 134
   : :|| :||| :||| :||| :||| :||| :||| :||| :||| :|||

RESULT 7
SYUB_RAT STANDARD; PRT; 137 AA.
AC O63754;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Beta-synuclein (Phosphonucleoprotein 14) (PNP 14).
GN SNCB.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=93019240; PubMed=1402909;
RA Tobé T., Nakajo S., Tanaka A., Mitoya A., Omata K., Nakaya K.,
RA Tomita M., Nakamura Y.;
RT "Cloning and characterization of the cDNA encoding a novel brain-
specific 14-kDa protein.";
RL J. Neurochem. 59:1624-1629(1992).
RN [2]
RP PHOSPHORYLATION BY CAM-KINASE II.
RX MEDLINE=94039126; PubMed=8223629;
RA Nakajo S., Tsukada K., Omata K., Nakamura Y., Nakaya K.;
RT "A new brain-specific 14-kDa protein is a phosphoprotein. Its complete
amino acid sequence and evidence for phosphorylation.";
RL Eur. J. Biochem. 217:1057-1063(1993).
CC -!- FUNCTION: MAY BE INVOLVED IN NEURONAL PLASTICITY.
CC -!- TISSUE SPECIFICITY: EXPRESSED SPECIFICALLY IN BRAIN.
CC -!- PTM: PHOSPHORYLATED. PHOSPHORYLATION BY G-PROTEIN COUPLED RECEPTOR
KINASES (GKR) IS MORE EFFICIENT THAN PHOSPHORYLATION BY CK1, CK2
AND CAM-KINASE II.
CC -!- SIMILARITY: BELONGS TO THE SYNUCLEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```



```

RT receptor kinases."
RL J. Biol. Chem. 275:26515-26522(2000).
CC -!- FUNCTION: PLAYS A ROLE IN NEUROFILAMENT NETWORK INTEGRITY. MAY BE
CC INVOLVED IN MODULATING AXONAL ARCHITECTURE DURING DEVELOPMENT AND
CC IN THE ADULT. IN VITRO, INCREASES THE SUSCEPTIBILITY OF
CC NEUROFILAMENT-H TO CALCIUM-DEPENDENT PROTEASES (BY SIMILARITY).
CC MAY ALSO FUNCTION IN MODULATING THE KERATIN NETWORK IN SKIN.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN BRAIN, PARTICULARLY IN THE
CC SUBSTANTIA NIGRA. ALSO EXPRESSED IN THE CORPUS CALLOSUM, HEART,
CC SKELETAL MUSCLE, OVARY, TESTIS, COLON AND SPLEEN. WEAK EXPRESSION
CC IN PANCREAS, KIDNEY AND LUNG.
CC -!- PTM: PHOSPHORYLATED. PHOSPHORYLATION BY GRK5 APPEARS TO OCCUR ON
CC RESIDUES DISTINCT FROM THE RESIDUE PHOSPHORYLATED BY OTHER
CC KINASES.
CC THE BREAST.
CC -!- DISEASE: BRAIN IRON ACCUMULATION TYPE 1 (NBIA1, ALSO CALLED
CC HALLERVORDEN-SPATZ SYNDROME), A RARE NEUROAXONAL DYSTROPHY, IS
CC HISTOLOGICALLY CHARACTERIZED BY AXONAL SPHEROIDS, IRON DEPOSITION,
CC LEWY BODY (LB)-LIKE INTRANEUROAL INCLUSIONS, GLIAL INCLUSIONS AND
CC NEUROFIBRILLARY TANGLES. SNCG IS FOUND IN SPHEROIDS BUT NOT IN
CC INCLUSIONS.
CC -!- SIMILARITY: BELONGS TO THE SYNUCLEIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF010126; AAB64109.1; -
CC EMBL; AF037207; AAC36586.1; -
CC EMBL; AF044311; AAC27738.1; -
CC EMBL; AF017256; AAC36550.1; -
CC EMBL; BC014098; AAL14098.1; -
CC Genew; HGNC:11141; SNCG.
CC MIM; 602998; -
CC InterPro; IPR001058; Synuclein.
CC Pfam; PF01387; Synuclein; 1.
CC PRINTS; PRO1211; Synuclein.
CC ProDom; PD010631; Synuclein; 1.
CC Repeat; Polymorphism; Phosphorylation.
CC DOMAIN 20 67
CC [EQ]-[GQ]-V-X(4).
CC REPEAT 20 30
CC 1.
CC 2.
CC REPEAT 31 41
CC REPEAT 42 56
CC REPEAT 57 67
CC MOD_RES 124 124
CC VARIANT 110 110
CC V->E.
CC /FTId=VAR_007455.
CC CONFLICT 13 13
CC CONFLICT 68 68
CC E -> K (IN REF. 1).
CC E -> K (IN REF. 1).
CC F19C046302DBE688 CRC64;
CC SEQUENCE 127 AA; 13301 MW; 13301 MW; F19C046302DBE688 CRC64;
CC
CC Query Match 44.6%; Score 313; DB 1; Length 127;
CC Best Local Similarity 62.8%; Pred. No. 1.5e-17;
CC Matches 71; Conservative 9; Mismatches 29; Indels 4; Gaps 2;
CC
CC QY 1 MDVFMKGLSKAKGVVAAAEKTKQGVAAEAKTKGVVGVGSKTKGVVGVGATVAEKTK 60
CC ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1 |||||
CC Db 1 MDVFMKGFSGIAKEGVVGVGAVKTKQGVTEAEKTKGVVGVGAKTKENVVQSVTSVAEKTK 60
CC
CC QY 61 EQVTVGGVAVTGVTVAAQKTVGAGSAGSAAATGFKKDKQGLKNGKCAP-QEGI 112
CC || ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1 |||||
CC Db 61 EQANAVSEAVSSVNTVATKTVEEAENIAVTSGVVRKEDL---RPSAPQQEGV 110
CC
CC RESULT 10
CC SYUG_MOUSE
CC ID SYUG_MOUSE STANDARD; PRT; 123 AA.

```

```

AC DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Gamma-synuclein (Persyn).
GN SNCG OR PERSYN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A., AND CHARACTERIZATION.
RP STRAIN=CD-1; TISSUE=Embryonic ganglion;
RX MEDLINE=99019751; Pubmed=9801372;
RA Buchman V.L., Hunter H.J., Pinon L.G., Thompson J., Privalova E.M.,
RA Ninkina N.N., Davies A.M.;
RT "Persyn, a member of the synuclein family, has a distinct pattern of
RT expression in the developing nervous system.";
RL J. Neurosci. 18:9335-9341(1998).
RN [2]
SEQUENCE FROM N.A.
RP STRAIN=CD-1;
RX MEDLINE=99162411; Pubmed=10051410;
RA Alimova-Kost M.V., Ninkina N.N., Imreh S., Gnuchev N.V., Adu J.,
RA Davies A.M., Buchman V.L.;
RT "Genomic structure and chromosomal localization of the mouse persyn
RT gene.";
RL Genomics 56:224-227(1999).
RN [3]
POSSIBLE FUNCTION IN SKIN.
RP MEDLINE=99126387; Pubmed=9925745;
RA Ninkina N.N., Privalova E.M., Pinon L.G., Davies A.M., Buchman V.L.;
RT "Developmentally regulated expression of persyn, a member of the
RT synuclein family, in skin.";
RL Exp. Cell Res. 246:308-311(1999).
CC -!- FUNCTION: PLAYS A ROLE IN NEUROFILAMENT NETWORK INTEGRITY. MAY BE
CC INVOLVED IN MODULATING AXONAL ARCHITECTURE DURING DEVELOPMENT AND
CC IN THE ADULT. IN VITRO, INCREASES THE SUSCEPTIBILITY OF
CC NEUROFILAMENT-H TO CALCIUM-DEPENDENT PROTEASES. MAY ALSO FUNCTION
CC IN MODULATING THE KERATIN NETWORK IN SKIN.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN BRAIN, PARTICULARLY IN THE
CC SUBSTANTIA NIGRA. ALSO EXPRESSED IN THE CORPUS CALLOSUM, HEART,
CC SKELETAL MUSCLE, OVARY, TESTIS, COLON AND SPLEEN. WEAK EXPRESSION
CC IN PANCREAS, KIDNEY AND LUNG. EXPRESSED PREDOMINANTLY IN THE CELL
CC BODIES AND AXONS OF PRIMARY SENSORY NEURONS, SYMPATHETIC NEURONS
CC AND MOTONEURONS.
CC -!- DEVELOPMENTAL STAGE: DEVELOPMENTALLY EXPRESSED IN PRIMARY SENSORY
CC NEURONS AND MOTONEURONS. IN TRIGEMINAL GANGLIA, EXPRESSION
CC INCREASES BETWEEN EMBRYONIC DAY 10 AND DAY 12. HIGH LEVELS ARE
CC MAINTAINED HERE THROUGHOUT LATER STAGES OF DEVELOPMENT AND IN
CC ADULTHOOD.
CC -!- PTM: PHOSPHORYLATED. PHOSPHORYLATION BY GRK5 APPEARS TO OCCUR ON
CC RESIDUES DISTINCT FROM THE RESIDUE PHOSPHORYLATED BY OTHER
CC KINASES (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE SYNUCLEIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF017255; AAC98893.1; -
CC EMBL; AF099986; AAD19899.1; -
CC EMBL; AF099984; AAD19899.1; JOINED.
CC EMBL; AF099985; AAD19899.1; JOINED.
CC SWISS-2DPAGE; Q920F7; MOUSE.
CC MGI; MGI:1298397; Sncc.
CC InterPro; IPR001058; Synuclein.
CC Pfam; PF01387; Synuclein; 1.

```

DR PRINTS; PRO1211; SYNUCLEIN.
 DR ProDom; PD010631; Synuclein; 1.
 KW Phosphorylation; Repeat.
 FT DOMAIN 20 67
 FT [EQ]-[GO]-V-X(4).
 FT REPEAT 20 30 1.
 FT REPEAT 31 41 2.
 FT REPEAT 42 56 3 (APPROXIMATE).
 FT REPEAT 57 67 4.
 FT MOD_RES 120 120
 FT PHOSPHORYLATION (BY CK2, CAMK2 AND GRK2)
 FT (BY SIMILARITY).
 FT SEQUENCE 123 AA; 13160 MW; DCB93C78F8071EE2 CRC64;
 SQ
 Query Match 44.5%; Score 312.5; DB 1; Length 123;
 Best Local Similarity 61.1%; Pred. No. 1.5e-17;
 Matches 69; Conservative 10; Mismatches 31; Indels 3; Gaps 1;
 QY 1 MDVFMKGLSKAKGEGVVAEAKTKGQVAAEAGTKGEGVLYVSGTKGEGVYVHGVATVAEKT 60
 DB 1 MDVFKGFSIAKEGVGAVGVEKTKGQVTEAAEKTGEGVMYVGTGKENVVQSVTSVAEKT 60
 QY 61 EQVTVNGGAVVTGTVAAQKTVEGAGSIAAATGFKKQDQ---GKNEGAPQE 110
 DB 61 EQANAVSEAVSVNTVANKTVEEAENIWTGTVRKEDLEPPAQDQAEKE 113
 RESULT 11
 SYUR_HUMAN
 ID SYUR_HUMAN STANDARD; PRT; 127 AA.
 AC Q9N250; Q9N187;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Synoretin (SR).
 OS Homo sapiens (Human), and
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606, 9913;
 RN [1]
 RP SEQUENCE FROM N.A., AND VARIANT PRO-106.
 RC SPECIES-Human, and Bovine; TISSUE-Retina;
 RX MEDLINE=99210388; PubMed=10192768;
 RA Surguchov A., Surgucheva I., Solessio E., Baehr W.;
 RT "Synoretin: a new protein belonging to the synuclein family.";
 RL Mol. Cell. Neurosci. 13:95-103(1999).
 RN [2]
 RP PHOSPHORYLATION BY G-PROTEIN COUPLED RECEPTOR KINASES.
 RX MEDLINE=20409007; PubMed=10852916;
 RA Pronin A.N., Morris A.J., Surguchov A., Benovic J.L.;
 RT "Synucleins are a novel class of substrates for G protein-coupled
 RT receptor kinases.";
 RL J. Biol. Chem. 275:26515-26522(2000).
 CC -1- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN RETINA
 CC (PREDOMINANTLY IN OUTER NUCLEAR LAYER, ALSO IN INNER SEGMENT OF
 CC PHOTORECEPTOR CELLS, SOME INDIVIDUAL CELLS LOCATED IN THE INNER
 CC NUCLEAR LAYER, INNER PLEXIFORM LAYER AND IN NERVE FIBER LAYER).
 CC ALSO FOUND IN BRAIN AND HEART.
 CC -1- PTM: PHOSPHORYLATED BY GRK2 AND GRK5.
 CC -1- SIMILARITY: BELONGS TO THE SYNUCLEIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF219257; AAF32342.1; -
 CC InterPro; IPR001058; Synuclein.
 CC Pfam; PF01387; Synuclein; 1.
 DR PRINTS; PR01211; SYNUCLEIN.

DR ProDom; PD010631; Synuclein; 1.
 KW Repeat; Polymorphism; Phosphorylation.
 FT DOMAIN 20 67
 FT [EQ]-[GO]-V-X(4).
 FT REPEAT 20 30 1.
 FT REPEAT 31 41 2.
 FT REPEAT 42 56 3 (APPROXIMATE).
 FT REPEAT 57 67 4.
 FT MOD_RES 124 124
 FT PHOSPHORYLATION (BY SIMILARITY).
 FT S -> P (HUMAN POLYMORPHISM).
 FT FTID=VAR_010900.
 FT SEQUENCE 127 AA; 13253 MW; 4271F507830C917D CRC64;
 SQ
 Query Match 43.8%; Score 307.5; DB 1; Length 127;
 Best Local Similarity 57.9%; Pred. No. 3.8e-17;
 Matches 70; Conservative 8; Mismatches 32; Indels 11; Gaps 1;
 QY 1 MDVFMKGLSKAKGEGVVAEAKTKGQVAAEAGTKGEGVLYVSGTKGEGVYVHGVATVAEKT 60
 DB 1 MDVFKGFSIAKEGVGAVGVEKTKGQVTEAAEKTGEGVMYVGTGKENVVQSVTSVAEKT 60
 QY 61 EQVTVNGGAVVTGTVAAQKTVEGAGSIAAATGFKKQDQ---DOLGKNEGAPQ 109
 DB 61 EQANAVSEAVSVNTVANKTVEEAENIWTGTVRKEDLEPPAQDQAEKE 120
 QY 110 E 110
 DB 121 E 121
 RESULT 12
 SYUG_RAT
 ID SYUG_RAT STANDARD; PRT; 123 AA.
 AC Q63544;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Gamma-synuclein (Parsyn) (Sensory neuron synuclein).
 GN SNCG.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Dorsal root ganglion;
 RX MEDLINE=95403363; PubMed=7673161;
 RA Akopian A.N., Wood N.;
 RT "Peripheral nervous system-specific genes identified by subtractive
 RT cDNA cloning.";
 RL J. Biol. Chem. 270:21264-21270(1995).
 RN [2]
 RP TISSUE SPECIFICITY.
 RX MEDLINE=99019751; PubMed=9801372;
 RA Buchman V.L., Hunter H.J., Pinon L.G., Thompson J., Privalova E.M.,
 RA Ninkina N.N., Davies A.M.;
 RT "Parsyn, a member of the synuclein family, has a distinct pattern of
 RT expression in the developing nervous system.";
 RL J. Neurosci. 18:9335-9341(1998).
 CC -1- FUNCTION: PLAYS A ROLE IN NEUROFILAMENT NETWORK INTEGRITY. MAY BE
 CC INVOLVED IN MODULATING AXONAL ARCHITECTURE DURING DEVELOPMENT AND
 CC IN THE ADULT. IN VITRO, INCREASES THE SUSCEPTIBILITY OF
 CC NEUROFILAMENT-H TO CALCIUM-DEPENDENT PROTEASES. MAY ALSO FUNCTION
 CC IN MODULATING THE KERATIN NETWORK IN SKIN.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- TISSUE SPECIFICITY: SPECIFICALLY EXPRESSED IN THE PERIPHERAL
 CC NERVOUS SYSTEM. HIGH EXPRESSION IN MOTONEURONS OF THE BRAINSTEM.
 CC ALSO FOUND IN NEURONS OF MANY OTHER BRAIN REGIONS INCLUDING THE
 CC CEREBELLAR CORTEX, THALAMUS, HYPOTHALAMUS AND CA1, CA2, CA3 AND CA4
 CC REGIONS OF THE HIPPOCAMPUS.
 CC -1- PTM: PHOSPHORYLATED. PHOSPHORYLATION BY GRK5 APPEARS TO OCCUR ON
 CC RESIDUES DISTINCT FROM THE RESIDUE PHOSPHORYLATED BY OTHER
 CC KINASES (BY SIMILARITY).

SQ SEQUENCE 464 AA; 48335 MW; BDD21CF95CCBAC7F CRC64;

Query Match 12.5%; Score 88; DB 1; Length 464;
Best Local Similarity 29.9%; Pred. No. 7;
Matches 43; Conservative 16; Mismatches 65; Indels 20; Gaps 6;

QY 8 LSKAKGVVAAAEKTKQGVAEAAA-----GKTKEGVLYVGSKTKEGVHGVATVAEKTKQE 62
: : : : : | | | | | : : : : : | | | | : :
Db 148 VQAEEAIVGHTEKATPMGAQEAETVGHIEKTTVG-----QAEEATV-GQAEEATVQGAEE 200

QY 63 VTWNGAVVTGTAVAKTVEGAGSIAATGFKVKDQLKNEE---GAPEF---GILEDM 116
| | | | | | | | | | : : : : : | | | | : :
Db 201 EATVGAQAEATVGAQAEATVGAQGE--ATVSHIERTTVGAQAEATVGAQAEATVGAQAEA 258

QY 117 PVPDPNEAYEMSEPGYQDYEPFA 140
| | | : : | | | | | | | | | | | | | | | |
Db 259 TVGQAEEATVGAQAEATVDOAEEA 282

RESULT 15
PURL_BACSU STANDARD; PRN: 742 AA.

AC P12042;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Phosphoribosylformylglycinamide synthase II (EC 6.3.5.3) (FGAM
synthase II).
GN PURL.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87250425; PubMed=3036807;
RZ Edbole D.J., Zaikin H.;
RT "Cloning and characterization of a 12-gene cluster from Bacillus
subtilis encoding nine enzymes for de novo purine nucleotide
synthesis.";
RL J. Biol. Chem. 262:8274-8287(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borcher S.,
Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
Guiseppi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
Hilbert H., Holsappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,
Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
Kobayashi Y., Koetter P., Konigstein G., Krogh S., Kumano M.,
Kurita K., Lapidus A., Lardinolis S., Lauber J., Lazarevic V.,
Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
Presecan E., Pujić P., Furnelle B., Rapoport G., Rey M., Reynolds S.,
Rieger M., Rivolta C., Roche B., Rose M., Sadale Y.,
Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
Sekiguchi J., Sekowska A., Serer S.J., Serron P., Shin B.S., Soldo B.,
Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
Takeuchi M., Tamakoshi A., Tanaka T., Terpestra P., Tognoni A.,
Tosato V., Uchiyama S., Vandenberg M., Vannier P., Vassarotti A.,
Viari A., Wambutt R., Wedler E., Wedler H., Weltzenegger T.,
Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RA "The complete genome sequence of the Gram-positive bacterium Bacillus

OX	NCBI_TaxID=9031;	
RP	[1]	
RN	SEQUENCE FROM N.A.	
RX	MEDLINE=20544973; PubMed=11091093;	
RA	Tluncova A.A., Anokhin K.V., Saba A.R., Schmidt O., Hanger D.P.,	
RA	Anderton B.H., Davies A.M., Ninkina N.N., Buchman V.L.;	
RT	"Chicken syncylineins: cloning and expression in the developing	
RT	embryo.";	
RL	Mech. Dev. 99:195-198(2000).	
DR	EMBL; AF253513; AAF67729.1; -;	
DR	InterPro; IPR001058; Synuclein.	
DR	Pfam; PF01387; Synuclein.1.	
DR	PRINTS; PR01211; SYNUCLEIN.	
DR	ProDom; PD010631; synuclein; 1.	
SQ	SEQUENCE 128 AA; 12966 MW; A5B94446053A2DA7 CRC64;	
Query Match 49.6%; Score 348.5; DB 13; Length 128;		
Best Local Similarity 60.4%; Pred. No. 7.9e-20;		
Matches 84; Conservative 7; Mismatches 35; Indels 13; Gaps		
QY	1 MDVFMKLSKAKGVVAAAEKTKQGVAAEAGKTKEGVLYVGSKTKEGVGHGVATVAEKT	60
Db	1 MDVFKKGFSTAKGVVAAAEKTKQGVTEAAEKTKEGVWVGTTKEGVQSVTSVAEKT	60
QY	61 EQVNTVGGAVVTGVTVAQKTVGAGSIAAATGVKKDQLGKNEGAPOEGILEMPVDP	120
Db	61 EQANVYGEAVVSVNTVANKTVEGAETIVATTGVKKEDL-----APQPPAAE	108
QY	121 DNEAYEMPSEGYQDYEP	139
Db	109 -GEAATPGSTEGGEGENE	126
RESULT 7		
Q96P61	PRELIMINARY; PRT; 127 AA.	
ID	Q96P61	
DC	Q96P61; PRELIMINARY; PRT; 127 AA.	
DT	01-DEC-2001 (TrEMBLrel. 19, Created)	
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)	
DT	01-MAR-2002 (TrEMBLrel. 20, Last annotation update)	
DE	Synuclein gamma.	
OS	Homo sapiens (human).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
OX	NCBI_TaxID=9606;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RA	Han C., Zhang B., Peng X., Yuan J., Qiang B.;	
RA	Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; AF411524; AAL05870.1; -;	
DR	InterPro; IPR001058; Synuclein.	
DR	Pfam; PF01387; Synuclein.1.	
DR	ProDom; PD010631; Synuclein; 1.	
SQ	SEQUENCE 127 AA; 13389 MW; D758B345DBFF7578 CRC64;	
Query Match 44.1%; Score 309.5; DB 4; Length 127;		
Best Local Similarity 62.7%; Pred. No. 8.2e-17;		
Matches 69; Conservative 9; Mismatches 29; Indels 3; Gaps		
QY	1 MDVFMKLSKAKGVVAAAEKTKQGVAAEAGKTKEGVLYVGSKTKEGVGHGVATVAEKT	60
Db	1 MDVFKKGFSTAKGVVDAVEKTKQGVTEAAEKTKEGVWVGAKTKENVVQSVTSVAEKT	60
QY	61 EQVNTVGGAVVTGVTVAQKTVGAGSIAAATGVKKDQLGKNEGAPOE	110
Db	61 EQANVSEAVSVSVNTVATKEAENIAVTGCVWRKEDL---RPSAPQ	107
RESULT 8		
Q9EQC3	PRELIMINARY; PRT; 54 AA.	
ID	Q9EQC3	
AC	Q9EQC3;	
DT	01-MAR-2001 (TrEMBLrel. 16, Created)	

```
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE Alpha-synuclein (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SVOLA;
RA Schluter O.M., Fornai F., Alessandri M.G., Geppert M., Jahn R.,
RA Sudhof T.C.;
RT "Deficiency of alpha-synuclein protects against MPTP-induced dopamine
RL depletion.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF277451; AAG44833.1; -.
DR InterPro; IPR001058; Synuclein.
DR Pfam; PF01387; Synuclein.1.
DR PRINTS; PD01211; SYNUCLEIN.
DR ProDom; PD010631; Synuclein; 1.
FT NON_TER 54 54
SQ SEQUENCE 54 AA; 5522 MW; C61A15427FE10CB9 CRC64;

Query Match 37.0%; Score 260; DB 11; Length 54;
Best Local Similarity 98.1%; Pred. No. 1.9e-13;
Matches 53; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDVFMKLSKAGGVVAAAEKTKQGVAAAGTKEGVLVVGSKTKEGVVGVAT 54
Db 1 MDVFMKLSKAGGVVAAAEKTKQGVAAAGTKEGVLVVGSKTKEGVVGVAT 54

RESULT 9
QY0044
ID QY0044 PRELIMINARY; PRT; 150 AA.
AC QY0044;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Durl35 (Fragment).
GN Durl35.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL-N2;
RA Cherkasova V.A., Egilmez N.K., Shmookler Reis R.J.;
RT "Diverse Caenorhabditis elegans genes that are upregulated in dauer
RL larvae are also induced by aging.";
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U52071; AAD11528.1; -.
DR InterPro; IPR001058; Synuclein.
DR PRINTS; PR01211; SYNUCLEIN.
FT NON_TER 1 1
SQ SEQUENCE 150 AA; 13614 MW; F14F153FECFE35C6 CRC64;

Query Match 17.5%; Score 123; DB 5; Length 150;
Best Local Similarity 33.6%; Pred. No. 0.027;
Matches 36; Conservative 12; Mismatches 39; Indels 20; Gaps 4;

QY 13 EGVVAAAEKTKQGVAAAGTKEGVLVVGSKTKEGVVGVATVAEKT---EQVTNVGGA 69
Db 10 DGAAVAKDAAVEGATTAGNKTVEGMRFVG-----ENVAHGAGKGVADGAKAGDTVVEGGKA 65

QY 70 VV----TGVTVAQKTVEGAGSIAAATGFVKDQLGKNEEGAPQEGI 112
Db 66 AAEKTESGIAFVGOKIAEGAGAVAGG-----AKAAGGAVVDGI 103

RESULT 10
Q95ZV0
ID Q95ZV0 PRELIMINARY; PRT; 750 AA.
AC Q95ZV0;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE F25H8.5c protein.
GN F25H8.5c.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Gajadsty S.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RL investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; Z69360; CAC42291.1; -.
DR InterPro; IPR004238; LEA.
DR Pfam; PF02987; LEA; 5.
SQ SEQUENCE 750 AA; 77336 MW; A68CA612D8345281 CRC64;

Query Match 17.4%; Score 122; DB 5; Length 750;
Best Local Similarity 32.7%; Pred. No. 0.24;
Matches 35; Conservative 13; Mismatches 39; Indels 20; Gaps 4;

QY 13 EGVVAAAEKTKQGVAAAGTKEGVLVVGSKTKEGVVGVATVAEKT---EQVTNVGGA 69
Db 610 DGAAVAKDAAVEGATTAGNKTVEGMRFVG-----ENVAHGAGKGVADGAKAGDTVVEGGKA 665

QY 70 VV----TGVTVAQKTVEGAGSIAAATGFVKDQLGKNEEGAPQEGI 112
Db 666 AAEKTESGIAFVGOKIAEGAGAVAGG-----AKAAGGAVVDGV 703

RESULT 11
Q19790
ID Q19790 PRELIMINARY; PRT; 798 AA.
AC Q19790;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE F25H8.5 protein.
GN F25H8.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Gajadsty S.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
CC -!- ALTERNATIVE PRODUCTS: TWO FORMS (A AND B) MAY BE PRODUCED BY
CC OF FORM A.
CC EMBL; Z69360; CAA93285.1; -.
DR EMBL; Z69360; CAA93286.1; -.
DR InterPro; IPR004238; LEA.
DR InterPro; IPR002965; P_rich_extensn.
DR Pfam; PF02987; LEA; 6.
DR PRINTS; PR01217; PRICHEXTENSIN.
KW Alternative splicing; Hypothetical protein.
FT VARSPLIC 387 387 N -> S (IN ISOFORM B).
FT VARSPLIC 388 798 MISSING (IN ISOFORM B).
SQ SEQUENCE 798 AA; 81507 MW; 4A61520E2313BFE2 CRC64;

Query Match 17.4%; Score 122; DB 5; Length 798;
Best Local Similarity 32.7%; Pred. No. 0.25;
```

Matches 35; Conservative 13; Mismatches 39; Indels 20; Gaps 4;

QY 13 EGVAATAAETKQGVAAAGKTEGVLYVGSKTKEGVVHGVATVAETK---EQVTVNGGA 69
 Db 610 DGAVAAKDAAVEGATTAGNKTGEMRFVG---ENVHAGAGKVAAGAKAAGDTVVGGKA 665

QY 70 VV-----TGVTVAAKTVGAGSTAAATGFKKDLGKNEEGAPQESG 112
 Db 666 AAEKTESGIAFGQKLAEGAGAVAGG-----AKAAGGAVVDGV 703

RESULT 12
 Q39846
 ID Q39846 PRELIMINARY; PRT; 643 AA.
 AC Q39846;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE LEA protein.
 OS Glycine max (Soybean).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
 OX NCBI_TaxID=3847;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SHI-SHI; TISSUE=COTYLEDON;
 RX MEDLINE=98418627; PubMed=9747855;
 RA Hsing Y.C., Tsou C.H., Hsu T.F., Chen Z.Y., Hsieh K.L., Hsieh J.S.,
 RA Chow T.Y.;
 RT "Tissue- and stage-specific expression of a soybean (Glycine max L.)
 RT seed-maturation, biotinylated protein.";
 RL Plant Mol. Biol. 38:481-490(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SHI-SHI; TISSUE=COTYLEDON;
 RA Shu T.F., Hsieh K.L., Hsing Y.I., Chen Z.Y., Chow T.Y.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U59626; AAC61783.1; -;
 SQ SEQUENCE 643 AA; 67878 MW; 0D18630D04196985 CRC64;

Query Match 15.0%; Score 105; DB 10; Length 643;
 Best Local Similarity 28.2%; Pred. No. 4;
 Matches 48; Conservative 10; Mismatches 68; Indels 44; Gaps 7;

QY 2 DVFMKGLSKAKGEGVAAAEKTKQGVAAAGKTKEG-----VLVYVGSK---TKEGVVHG--- 51
 Db 289 DTISSAAKTASEKTPVAAEKADYTLQAAEKAKSAGGTTASYVGBKVAQKDVAVESGKS 348

QY 52 ----VATVAEKTEQVTVNGGA-----VVTGTVAAQKTVGAG----- 86
 Db 349 AAGYAAKVAADLRDKATATVGAHAFSAEKTVEGTAKAAH-VVEGAAGYAGHKAELASM 407

QY 87 -----STAAATGFKVKKDLGKNEEGAPQEGILEMPVDPDNEAYEMPSE 130
 Db 408 SAGAVKGLAASAGETAKEYTAKKKEAQRE--LEAKKPSQPOEAERPSE 455

RESULT 13
 Q39967
 ID Q39967 PRELIMINARY; PRT; 317 AA.
 AC Q39967;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Group 3 late embryogenesis abundant protein (Fragment).
 OS Triticum aestivum (Wheat).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
 OC Triticeae; Triticum.
 OX NCBI_TaxID=4565;
 RN [1]
 RP SEQUENCE FROM N.A.

TISSUE=EMBRYONIC AXES;
 RC MEDLINE=93222487; PubMed=8467082;
 RA Walker-Simmons M.K., Curry J.;
 RT "Unusual sequence of group 3 LEA (II) mRNA inducible by dehydration
 RT stress in wheat.";
 RL Plant Mol. Biol. 21:907-912(1993).
 DR EMBL; M72395; AAA34267.1; -;
 DR InterPro; IPR004238; LEA.
 DR Pfam; PF02987; LEA; 3.
 FT NON_TER 1
 SQ SEQUENCE 317 AA; 33367 MW; 75BB0F431220160C CRC64;

Query Match 14.2%; Score 100; DB 10; Length 317;
 Best Local Similarity 29.6%; Pred. No. 4.1;
 Matches 53; Conservative 10; Mismatches 62; Indels 54; Gaps 8;

QY 6 KGLSKAKGEGVAAAEKTKQGVAAE-----AAGTKTEGVLYVGSKTKEGVVHGVATVAE 57
 Db 9 EGKEGASEMADAAAGTKDAAAEXTREMAADAAAGTKETKDAAVEKTRE-MADTAATKAA 67

QY 58 KTK-----EQVT-----NVGGAVVT-----GVTAAQKTVGA--- 85
 Db 68 ETKDAAAEKASGAGEMVTEKARSADAAADKASGAETVTEKAKGAKDAALDTAEGAKEY 127

QY 86 -----GSIAATGFKVKKDLGKNEEGAPQEGILEMPVDPDNEAYEMPSEGVQDVEP 138
 Db 128 MVDKKEDARRALAGSV-KDSKGETNESACQOG-----QDVRRAAEKAEARQTHOP 179

RESULT 14
 Q8SVY9
 ID Q8SVY9 PRELIMINARY; PRT; 166 AA.
 AC Q8SVY9;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Hypothetical protein ECU03_1610.
 GN ECU03_1610.
 OS Encephalitozoon cuniculi.
 OC Eukaryota; Microsporidia; Unikaryonidae; Encephalitozoon.
 OX NCBI_TaxID=6035;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GB-M1;
 RA Genoscope;
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GB-M1;
 RX MEDLINE=21576510; PubMed=11719806;
 RA Katinka M.D., Duprat S., Cornilliot E., Metenier G., Thomarat F.,
 RA Prentiss G., Barbe V., Peyretailade E., Brottier P., Wincker P.,
 RA Delbac F., El Alaoui H., Peyret P., Saurin W., Gouy M.,
 RA Welssbach J., Vivares C.P.;
 RT "Genome sequence and gene compaction of the eukaryote parasite
 RT Encephalitozoon cuniculi.";
 RL Nature 414:450-453(2001).
 RL EMBL; AL590443; CAD26304.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 166 AA; 18030 MW; C1E7E4EA2E31D181 CRC64;

Query Match 14.1%; Score 99; DB 5; Length 166;
 Best Local Similarity 30.5%; Pred. No. 2.2;
 Matches 39; Conservative 11; Mismatches 62; Indels 16; Gaps 4;

QY 10 KAKGEGVAAAEKTKQGVAAAGKTKEGVLVYVGSKTKEGVVHGVATVAEKTKEQVTVNGGA 69
 Db 21 KANDKACEARDKTKETAGSAKDKTKETAGSAKDKTKETAGSAKDKTKETAGS 72

QY 70 VVTGTVAAQ-----KTVGAGASTAAATGFKVKKDLGKNEEGAPQEGILEMPVDPD-NEA 124
 Db 73 AKDKTKETAGSAKDKTKETAGSAKDKTKETAGSAKDKTKETAGSAKDKTKETAGS 129

QY 125 YEMPSEEG 132
 Db 130 YENIKEGG 137

RESULT 15

Q971D3 PRELIMINARY; PRT; 380 AA.
 AC Q971D3;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Hypothetical protein ST1420.
 GN ST1420.
 OS Sulfolobus tokodaii.
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
 OC Sulfolobus.
 OX NCBI_TaxID=111955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JCM 10545 / 7;
 RX PubMed=11572479;
 RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
 RA Sekine M., Baba S.-I., Ankai A., Kosugi H., Hosoyama A., Fukui S.,
 RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
 RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
 RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
 RA Oshima T., Kikuchi H.;
 RT "Complete genome sequence of an aerobic thermoacidophilic
 RT Crenarchaeon, Sulfolobus tokodaii strain7.";
 RL DNA Res. 8:123-140(2001).
 DR EMBL; AP000986; BAB66487.1; .
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 380 AA; 42015 MW; 859AC76EE032D719 CRC64;

Query Match 14.1%; Score 99; DB 17; Length 380;
 Best Local Similarity 28.6%; Pred. No. 6.1;
 Matches 42; Conservative 19; Mismatches 66; Indels 20; Gaps 6;
 QY 2 DVFEMKGLS-----KAREGVVAAREKTKQGVAAEAGKTK-EGVLYVGSKTKEGVVHGV 52
 Db 197 DVFELGIPVGDLSGLPLVASRFLNISEKWSPSREMTAGETEFEGRLIVLKA-EGPMATV 255
 QY 53 ATVAERTKEQVTWVGAVTGTVAQKTVEG--AGSIAAATGFVKKQDLGKNEGAPQE 110
 Db 256 GTFGEAVENVVNLGCKVSRITVDAAAKLEGGQTGCSIAEGTGVAMGDP-GPEKTAIERV 314
 QY 111 GILEDPVDP-----DNEAYEMPSE 130
 Db 315 AVRYNIPIDALIVKMSMEEAITEMPRE 341

Search completed: January 28, 2003, 14:18:18
 Job time : 9.07457 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 13:58:08 ; Search time 7.80115 Seconds
(without alignments)
2391.326 Million cell updates/sec

Title: US-09-904-987-4

Perfect score: 702

Sequence: 1 MDVFMKGLSKAKGVVAAAE.....DNEAYEMPSEGVQDYEPFA 140

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002:*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	702	100.0	140	16	AA70127
2	702	100.0	140	20	AA70727
3	702	100.0	140	20	AAW88131
4	702	100.0	140	21	AA702093
5	702	100.0	140	21	AA751529
6	702	100.0	140	23	ABB1545
7	697	99.3	140	21	AA702094
8	695	99.0	140	21	AA702090
9	693	98.7	140	21	AA702095
10	692	98.6	140	20	AAW88132

11	692	98.6	140	21	AA702091	Artificial alpha-s
12	692	98.6	140	21	AA702096	Artificial alpha-sy
13	684	97.4	140	21	AA702092	Artificial alpha-s
14	670	95.4	140	20	AA707274	Rat synuclein alte
15	663	94.4	140	21	AA702220	NACP/alpha-synucle
16	595.5	84.8	143	20	AA707280	Zebrafinch (sic) s
17	485	69.1	149	20	AA707275	Rat synuclein alte
18	427.5	60.9	133	20	AA707279	Bovine synuclein.
19	409	58.3	134	20	AA707278	Human beta-synucle
20	313	44.6	127	20	AA707271	Human beta-synucle
21	312.5	44.5	123	20	AA707272	Mouse persyn prote
22	311.5	44.4	163	21	AAW58881	Breast and ovarian
23	311.5	44.4	164	22	AAO13847	Human polypeptide
24	307.5	43.8	127	19	AAW63123	Human breast cancer
25	307.5	43.8	127	21	AA707779	Human HBGA67A pro
26	273.5	39.0	123	20	AA707281	Rat synuclein-like
27	261	37.2	113	20	AA707273	Torpedo californic
28	251	35.8	51	21	AAW02343	Human secreted pro
29	226	32.2	47	22	ABB43301	Peptide #10807 enc
30	226	32.2	47	22	AAW64209	Human brain expres
31	226	32.2	47	22	AAW77035	Human bone marrow
32	216	30.8	59	21	AAW02228	Human secreted pro
33	199	28.3	42	20	AA707276	Rat synuclein alte
34	176	25.1	37	22	AAE08570	Human alpha-synuc
35	167	23.8	37	22	AAE08564	Human Beta-synuc
36	166	23.6	35	22	AAW91804	Anyloid beta-prote
37	166	23.6	35	22	AAW91831	Anyloid beta-prote
38	163	23.2	35	16	AA70128	Novel amyloid comp
39	151	21.5	28	22	AAW65485	Human brain expres
40	151	21.5	28	22	AAW78149	Human bone marrow
41	151	21.5	28	23	ABG47162	Human peptide enco
42	143	20.4	30	22	AAW41757	Peptide #9283 enco
43	143	20.4	30	22	AAW75444	Human bone marrow
44	143	20.4	30	22	AAW35555	Peptide #9592 enco
45	142	20.2	37	22	AAE08563	Human Beta-synucle

ALIGNMENTS

RESULT 1
AA70127
ID AA70127 standard; Protein; 140 AA.
XX
AC AA70127;
XX
DT 05-OCT-1995 (first entry)
XX
DE Precursor of novel amyloid component (NACP).
XX
KW Amyloid component; NAC; Alzheimer's disease.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 61..80
FT Peptide /label= x peptide 81..95
FT Peptide /label= y peptide
XX
PN WO9506407-A.
XX
PD 09-MAR-1995.
XX
PF 29-AUG-1994; 94WO-US09789.
XX
PR 30-AUG-1993; 93US-0114393.
XX
PA (SAIT/) SAITOH T.
XX
PI (REGC) UNIV CALIFORNIA.
XX
Masliah E;

DR WPI; 1995-115208/15.
 DR N-PSDB; AAQ83201.
 XX
 PT Novel component of amyloid in Alzheimer's disease - used in
 PT methods to diagnose and treat neuronal abnormalities in
 PT particular the deposition of amyloid plaques characteristic of AD
 XX
 PS Claim 66; Fig 2A-1 - 2A-2; 128pp; English.
 XX
 CC AAQ83201 encodes a precursor of the novel amyloid component (NACP)
 CC (AAQ70127) with an apparent mol. mass of 19,000 Da. NACP is the
 CC precursor of NAC (AAQ70128). In vivo, NACP is recovered in the
 CC cytosolic fraction of human brain homogenate. NACP has seven
 CC repeated KTRGV AA motifs, but no signal peptide sequence nor
 CC N-linked glycosylation sites. NAC is located in the most hydrophobic
 CC portion of NACP. NAC has a mol. wt. of approx. 3,500 Mr. within NAC,
 CC two new amyloid sequences have been identified known as the 'X' and
 CC 'Y' peptides which are encoded contiguously in the most hydrophobic
 CC domain (see AAQ70127 FT). The X and Y peptides are claimed (claims 68
 CC and 69 respectively).
 XX
 SQ Sequence 140 AA;

Query Match 100.0%; Score 702; DB 16; Length 140;
 Best Local Similarity 100.0%; Pred. No. 4e-60;
 Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDVFMKGLSKAKGVAAGKTKQGVAAAGTKGVLVVGSKTKGVVHGVATVAEKT 60
 DB 1 MDVFMKGLSKAKGVAAGKTKQGVAAAGTKGVLVVGSKTKGVVHGVATVAEKT 60
 QY 61 EQVTNVGGAVVTGVTVAQAQTVGAGSIAAATGFVKKQDLGKNEEGAPQEGILEDMPVDP 120
 DB 61 EQVTNVGGAVVTGVTVAQAQTVGAGSIAAATGFVKKQDLGKNEEGAPQEGILEDMPVDP 120
 QY 121 DNEAYEMPSEEGYQDYEP 140
 DB 121 DNEAYEMPSEEGYQDYEP 140

RESULT 2
 AAQ70277
 ID AAQ70277 standard; Protein: 140 AA.

AC AAQ70277;
 XX
 XX 06-JUL-1999 (first entry)
 DE Human alpha-synuclein.
 XX
 KW Human; mouse; synuclein; persyn; diagnosis; neurodegenerative disorder;
 KW cancer; breast; skin; intermediate filament damage.
 XX
 OS Homo sapiens.
 XX
 PN EP908727-A1.
 XX
 XX 14-APR-1999.
 XX
 XX 21-SEP-1998; 98EP-0307628.
 XX
 XX 19-SEP-1997; 97GB-0019879.
 XX
 PA (NEUR-) EUROPA LTD.
 PA (UYSA-) UNIV ST ANDREWS.
 XX
 XX WPI; 1999-217169/19.
 DR
 XX New synuclein protein (persyn) and gene, useful in assays for
 PT screening, diagnosing or monitoring cancer, neurodegenerative
 PT disorders or skin disorders
 XX
 PS Disclosure; Page 22-23; 39pp; English.

XX This sequence represents the sequence of the human alpha-synuclein.
 CC The invention relates to novel human and mouse synuclein family members
 CC designated persyn (AAQ70271 and AAQ70172). The persyn sequence is
 CC useful for screening, diagnosing or monitoring cancer (especially breast
 CC or skin cancer), neurodegenerative disorders or skin disorders and for
 CC identifying cells having intermediate filament damage.
 XX
 SQ Sequence 140 AA;

Query Match 100.0%; Score 702; DB 20; Length 140;
 Best Local Similarity 100.0%; Pred. No. 4e-60;
 Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDVFMKGLSKAKGVAAGKTKQGVAAAGTKGVLVVGSKTKGVVHGVATVAEKT 60
 DB 1 MDVFMKGLSKAKGVAAGKTKQGVAAAGTKGVLVVGSKTKGVVHGVATVAEKT 60
 QY 61 EQVTNVGGAVVTGVTVAQAQTVGAGSIAAATGFVKKQDLGKNEEGAPQEGILEDMPVDP 120
 DB 61 EQVTNVGGAVVTGVTVAQAQTVGAGSIAAATGFVKKQDLGKNEEGAPQEGILEDMPVDP 120
 QY 121 DNEAYEMPSEEGYQDYEP 140
 DB 121 DNEAYEMPSEEGYQDYEP 140

RESULT 3
 AAQ88131
 ID AAQ88131 standard; Protein: 140 AA.

XX AAQ88131;
 XX 24-MAY-1999 (first entry)
 XX Human alpha synuclein protein.
 DE Synuclein; Parkinson disease; drug screening; diagnosis.

OS Homo sapiens.
 PN WO9859050-A1.
 XX 30-DEC-1998.
 XX 25-JUN-1998; 98WO-US13071.
 XX 25-JUN-1997; 97US-0050684.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX Duvoisin RC, Johnson WG, Lavedan C, Leroy E, Nussbaum RL;
 PI Polymeropoulos MH;
 XX WPI; 1999-105624/09.
 XX New gene mutation associated with Parkinson's disease - comprising a
 PT mutation in the synuclein gene, used to develop products for
 PT detecting a predisposition to or treating Parkinson's disease
 XX
 PS Disclosure; Page 50-51; 96pp; English.

XX A mutated human synuclein protein or homologue, associated with
 CC predisposition to Parkinson's disease (PD). The products can be used for
 CC testing for predisposition to PD. They can also be used for studying the
 CC pathophysiology of PD. They can also be used for identifying compounds
 CC for inhibiting the self-aggregation of mutant proteins that is thought
 CC to lead to PD. The products can also be used in methods directed at the
 CC correction or suppression of PD.
 XX
 SQ Sequence 140 AA;

Query Match 100.0%; Score 702; DB 20; Length 140;


```

CC 32.
XX SQ Sequence 140 AA;
Query Match 99.3%; Score 697; DB 21; Length 140;
Best Local Similarity 99.3%; Pred. No. 1.2e-59;
Matches 139; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MDVFMKGLSKAKGVAAGAAKTKQGVAAEAGKTKGVLVYVGSKTKEGVVHGVAATVAEKT 60
DB 1 MDVFMKGLSKAKGVAAGAAKTKQGVAAEAGKTKGVLVYVGSKTKEGVVHGVAATVAEKT 60
QY 61 EQVTNVGGAVVTGVTAVAQKTVEGAGSIAAATGFKVKKDQLGKNEGAPOEGILEDMPVDP 120
DB 61 EQVTNVGGAVVTGVTAVAQKTVEGAGSIAAATGFKVKKDQLGKNEGAPOEGILEDMPVDP 120
QY 121 DNEAYEMPSEEGYQDYEP 140
DB 121 DNEAYEMPSEEGYQDYEP 140
RESULT 8
AAY92090
ID AAY92090 standard; Protein; 140 AA.
XX AC AAY92090;
XX DT 01-AUG-2000 (first entry)
XX DE Artificial alpha-synuclein mutant B83Q+A90V.
XX KW Artificial; alpha-synuclein mutant; accelerated aggregation;
XX KW animal model; inhibitor; Parkinson's Disease.
XX OS Homo sapiens.
XX OS Synthetic.
XX PN WO200018917-A2.
XX PD 06-APR-2000.
XX PF 27-SEP-1999; 99WO-US22302.
XX PR 25-SEP-1998; 98US-0101862.
XX PR 24-SEP-1999; 99US-0405035.
XX PA (AMGE-) AMGEN INC.
XX PI Biere AL, Citron M;
XX DR WPI; 2000-293156/25.
XX PT New artificial mutant alpha-synuclein polypeptides are useful to
PT generate an animal model for Parkinson's disease and aggregate faster
PT than the naturally occurring alpha-synuclein and alpha-synuclein
PT mutants
XX PS Claim 3; Page 32-33; 34pp; English.
XX CC AAY92090-92 are artificial alpha-synuclein mutant polypeptides which
CC demonstrate accelerated aggregation compared to naturally occurring
CC alpha-synuclein mutant polypeptides. Aggregate formation was enhanced
CC for the E83Q+A90V and H50Y+A53T mutants, even when compared to the A53T
CC mutant. The aggregate formation was significantly enhanced for the
CC H50T+A53T+A76T mutant. An in vitro aggregation assay for evaluating
CC alpha-synuclein mutants which demonstrate accelerated aggregation is
CC also claimed. Additionally, a transgenic non-human mammal which has
CC been genetically engineered to express the mutant alpha-synuclein
CC polypeptide in the brain and an in vitro aggregation assay for high
CC throughput screening of alpha-synuclein nucleation inhibitors are
CC claimed. Alpha-synuclein inhibitors are used to treat, prevent or
CC ameliorate Parkinson's Disease or related disorders (claimed). The
CC transgenic animal is used as an animal model for Parkinson's Disease.

```

```

XX SQ Sequence 140 AA;
Query Match 99.0%; Score 695; DB 21; Length 140;
Best Local Similarity 98.6%; Pred. No. 1.9e-59;
Matches 138; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 MDVFMKGLSKAKGVAAGAAKTKQGVAAEAGKTKGVLVYVGSKTKEGVVHGVAATVAEKT 60
DB 1 MDVFMKGLSKAKGVAAGAAKTKQGVAAEAGKTKGVLVYVGSKTKEGVVHGVAATVAEKT 60
QY 61 EQVTNVGGAVVTGVTAVAQKTVEGAGSIAAATGFKVKKDQLGKNEGAPOEGILEDMPVDP 120
DB 61 EQVTNVGGAVVTGVTAVAQKTVEGAGSIAAATGFKVKKDQLGKNEGAPOEGILEDMPVDP 120
QY 121 DNEAYEMPSEEGYQDYEP 140
DB 121 DNEAYEMPSEEGYQDYEP 140
RESULT 9
AAY92095
ID AAY92095 standard; Protein; 140 AA.
XX AC AAY92095;
XX DT 01-AUG-2000 (first entry)
XX DE Artificial alpha-synuclein mutant A53T+A30P.
XX KW Artificial; alpha-synuclein mutant; accelerated aggregation;
XX KW animal model; inhibitor; Parkinson's Disease.
XX OS Homo sapiens.
XX OS Synthetic.
XX PN WO200018917-A2.
XX PD 06-APR-2000.
XX PF 27-SEP-1999; 99WO-US22302.
XX PR 25-SEP-1998; 98US-0101862.
XX PR 24-SEP-1999; 99US-0405035.
XX PA (AMGE-) AMGEN INC.
XX PI Biere AL, Citron M;
XX DR WPI; 2000-293156/25.
XX PT New artificial mutant alpha-synuclein polypeptides are useful to
PT generate an animal model for Parkinson's disease and aggregate faster
PT than the naturally occurring alpha-synuclein and alpha-synuclein
PT mutants
XX PS Example 1; Page -; 34pp; English.
XX CC This is an artificial alpha-synuclein double mutant polypeptide
CC A53T+A30P. AAY92090-92 are artificial alpha-synuclein mutant polypeptides
CC which demonstrate accelerated aggregation compared to naturally
CC occurring alpha-synuclein mutant polypeptides. Aggregate formation was
CC enhanced for the E83Q+A90V and H50Y+A53T mutants, even when compared to
CC the A53T mutant. The aggregate formation was significantly enhanced for
CC the H50T+A53T+A76T mutant. An in vitro aggregation assay for evaluating
CC alpha-synuclein mutants which demonstrate accelerated aggregation is
CC also claimed. Additionally, a transgenic non-human mammal which has
CC been genetically engineered to express the mutant alpha-synuclein
CC polypeptide in the brain and an in vitro aggregation assay for high
CC throughput screening of alpha-synuclein nucleation inhibitors are
CC claimed. Alpha-synuclein inhibitors are used to treat, prevent or
CC ameliorate Parkinson's Disease or related disorders (claimed). The
CC transgenic animal is used as an animal model for Parkinson's Disease.

```

CC This sequence is not given in the specification, it was constructed
 CC using the wild type sequence, which is given in the specification on page
 CC 32.

XX
 XX
 SQ Sequence 140 AA; Query Match 98.7%; Score 693; DB 21; Length 140;
 Best Local Similarity 98.6%; Pred. No. 2.9e-59;
 Matches 138; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MDVFMKGLSKAKGEGVVAAGKTKQGVAAAGKTKGVLVVGSKTKGEGVHGVTVAEKT 60
 DB 1 MDVFMKGLSKAKGEGVVAAGKTKQGVAAAGKTKGVLVVGSKTKGEGVHGVTVAEKT 60
 QY 61 EQVTNVGGAVVTGVTVAQAQTVGAGSIAAATGFVKKDLGKNEGAPQEGILEDMPVDP 120
 DB 61 EQVTNVGGAVVTGVTVAQAQTVGAGSIAAATGFVKKDLGKNEGAPQEGILEDMPVDP 120

QY 121 DNEAYEMPSEEGYQDYEP 140
 DB 121 DNEAYEMPSEEGYQDYEP 140

QY 121 DNEAYEMPSEEGYQDYEP 140
 DB 121 DNEAYEMPSEEGYQDYEP 140

QY 121 DNEAYEMPSEEGYQDYEP 140
 DB 121 DNEAYEMPSEEGYQDYEP 140

RESULT 10
 AAW88132
 ID AAW88132 standard; Protein; 140 AA.

XX
 AC AAW88132;
 XX
 XX 24-MAY-1999 (first entry)

XX Human alpha synuclein protein mutant.
 XX Synuclein; Parkinson disease; drug screening; diagnosis.

XX Homo sapiens.
 OS Synthetic.

XX
 FH Key Location/Qualifiers
 FT Modified-site 53
 FT /label= A53-D53

XX W09859050-A1.

XX 30-DEC-1998.

XX 25-JUN-1998; 98WO-US13071.

XX 25-JUN-1997; 97US-0050684.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Duvoisin RC, Johnson WG, Lavedan C, Leroy E, Nussbaum RL;
 PI Polymeropoulos MH;

XX WPI; 1999-105624/09.

XX New gene mutation associated with Parkinson's disease - comprising a
 PT mutation in the synuclein gene, used to develop products for
 PT detecting a predisposition to or treating Parkinson's disease

XX Disclosure; Page-; 96pp; English.

XX A mutated human synuclein protein or homologue, associated with
 CC predisposition to Parkinson's disease (PD). The products can be used for
 CC testing for predisposition to PD. They can also be used for studying the
 CC pathophysiology of PD. They can also be used for identifying compounds
 CC for inhibiting the self-aggregation of mutant proteins that is thought
 CC to lead to PD. The products can also be used in methods directed at the
 CC correction or suppression of PD.
 CC Note: this sequence was created from information provided in the
 CC specification by the indexer.

XX

SQ Sequence 140 AA;

Query Match 98.6%; Score 692; DB 20; Length 140;
 Best Local Similarity 98.6%; Pred. No. 3.7e-59;
 Matches 138; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MDVFMKGLSKAKGEGVVAAGKTKQGVAAAGKTKGVLVVGSKTKGEGVHGVTVAEKT 60
 DB 1 MDVFMKGLSKAKGEGVVAAGKTKQGVAAAGKTKGVLVVGSKTKGEGVHGVTVAEKT 60

QY 61 EQVTNVGGAVVTGVTVAQAQTVGAGSIAAATGFVKKDLGKNEGAPQEGILEDMPVDP 120
 DB 61 EQVTNVGGAVVTGVTVAQAQTVGAGSIAAATGFVKKDLGKNEGAPQEGILEDMPVDP 120

QY 121 DNEAYEMPSEEGYQDYEP 140
 DB 121 DNEAYEMPSEEGYQDYEP 140

QY 121 DNEAYEMPSEEGYQDYEP 140
 DB 121 DNEAYEMPSEEGYQDYEP 140

RESULT 11
 AAY92091
 ID AAY92091 standard; Protein; 140 AA.

XX
 AC AAY92091;

XX
 DT 01-AUG-2000 (first entry)

XX Artificial alpha-synuclein mutant A53T.

XX Artificial; alpha-synuclein mutant; accelerated aggregation;
 KW animal model; inhibitor; Parkinson's Disease.

XX Homo sapiens.
 OS Synthetic.

XX W0200018917-A2.

XX 06-APR-2000.

XX 27-SEP-1999; 99WO-US22302.

XX 25-SEP-1998; 98US-0101862.

XX 24-SEP-1999; 99US-0405035.

XX (AMGE-) AMGEN INC.

XX Biere AL, Citron M;

XX WPI; 2000-293156/25.

XX New artificial mutant alpha-synuclein polypeptides are useful to
 PT generate an animal model for Parkinson's disease and aggregate faster
 PT than the naturally occurring alpha-synuclein and alpha-synuclein
 PT mutants

XX Claim 3; Page 33; 34pp; English.

XX AAY92090-92 are artificial alpha-synuclein mutant polypeptides which
 CC demonstrate accelerated aggregation compared to naturally occurring
 CC alpha-synuclein mutant polypeptides. Aggregate formation was enhanced
 CC for the E83Q-A90V and H50Y-A53T mutants, even when compared to the A53T
 CC mutant. The aggregate formation was significantly enhanced for the
 CC H50T-A53T-A76T mutant. An in vitro aggregation assay for evaluating
 CC alpha-synuclein mutants which demonstrate accelerated aggregation is
 CC also claimed. Additionally, a transgenic non-human mammal which has
 CC been genetically engineered to express the mutant alpha-synuclein
 CC polypeptide in the brain and an in vitro aggregation assay for high
 CC throughput screening of alpha-synuclein nucleation inhibitors are
 CC claimed. Alpha-synuclein inhibitors are used to treat, prevent or
 CC ameliorate Parkinson's Disease or related disorders (Claimed). The
 CC transgenic animal is used as an animal model for Parkinson's Disease.

XX Sequence 140 AA;


```

XX SQ Sequence 140 AA;
Query Match 97.4%; Score 684; DB 21; Length 140;
Best Local Similarity 97.9%; Pred. No. 2.2e-58;
Matches 137; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MDVFMKGLSKAKGEGVAAAEKTKGVAAAGKTKGVLVYVGSKTKGCVHGVATVAEKK 60
DB 1 MDVFMKGLSKAKGEGVAAAEKTKGVAAAGKTKGVLVYVGSKTKGCVHGVATVAEKK 60

QY 61 EQVTNNGGAVVTGVTAVAQKTVEGAGSIAAATGFKVKKDQLGKNEEGAPQEGILEDMPVDP 120
DB 61 EQVTNNGGAVVTGVTAVAQKTVEGAGSIAAATGFKVKKDQLGKNEEGAPQEGILEDMPVDP 120

QY 121 DNEAYEMPSEEGYQDYEP 140
DB 121 DNEAYEMPSEEGYQDYEP 140

RESULT 14
AAY07274
ID AAY07274 standard; Protein; 140 AA.
XX AC AAY07274;
XX FH Location/Qualifiers
XX FT 73..83
XX FT /note= "missing from beta-synuclein"
XX PN WO200020020-A2.
XX PD 13-APR-2000.
XX PF 06-OCT-1999; 99WO-US23134.
XX PR 06-OCT-1998; 98US-0103310.
XX PA (REGC ) UNIV CALIFORNIA.
XX PI Masliah E;
XX DR WPI; 2000-303637/26.
XX PT Screening treatments for neurodegenerative disease, such as Alzheimer's
XX PT disease, comprises stimulating oxidative stress and measuring and
XX PT aggregation of non-amyloid beta component /alpha-synuclein in neuronal
XX PT cells
XX PS Example 4; Page 47; 47pp; English.
XX CC The invention concerns methods for screening treatments for, and
XX CC treatment of, neurodegenerative disease. The methods comprise measuring
XX CC aggregation of neurons of NACP/alpha-synuclein and stimulating expression
XX CC of a non-amyloidogenic protein in order to reduce the level of
XX CC aggregation. In the screening method, oxidative stress is stimulated in
XX CC the neuronal cells by introduction of metal ions and hydrogen peroxide.
XX CC The methods are useful for treating and screening treatments for
XX CC neurodegenerative disease consisting of Alzheimer's disease, Parkinson's
XX CC disease and Lewy body disease.
XX SQ Sequence 140 AA;
Query Match 95.4%; Score 670; DB 20; Length 140;
Best Local Similarity 95.8%; Pred. No. 4.8e-57;
Matches 133; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 MDVFMKGLSKAKGEGVAAAEKTKGVAAAGKTKGVLVYVGSKTKGCVHGVATVAEKK 60
DB 1 MDVFMKGLSKAKGEGVAAAEKTKGVAAAGKTKGVLVYVGSKTKGCVHGVATVAEKK 60

QY 61 EQVTNNGGAVVTGVTAVAQKTVEGAGSIAAATGFKVKKDQLGKNEEGAPQEGILEDMPVDP 120
DB 61 EQVTNNGGAVVTGVTAVAQKTVEGAGSIAAATGFKVKKDQLGKNEEGAPQEGILEDMPVDP 120

```

```

DB 61 EQVTNNGGAVVTGVTAVAQKTVEGAGSIAAATGFKVKKDQMGKGEYPOEGILEDMPVDP 120
QY 121 DNEAYEMPSEEGYQDYEP 140
DB 121 DNEAYEMPSEEGYQDYEP 140

RESULT 15
ID AAY92220
XX AC AAY92220 standard; Protein; 140 AA.
XX AC AAY92220;
XX DT 10-AUG-2000 (first entry)
XX DE NACP/alpha-synuclein.
XX KW NACP/alpha-synuclein; beta-synuclein; beta-synuclein delta 1; mutant;
XX KW primer; neurodegenerative; Alzheimers disease; Parkinsons disease;
XX KW Lewy body disease; transgene.
XX OS Mus musculus.
XX FH Location/Qualifiers
XX FT 73..83
XX FT /note= "missing from beta-synuclein"
XX PN WO200020020-A2.
XX PD 13-APR-2000.
XX PF 06-OCT-1999; 99WO-US23134.
XX PR 06-OCT-1998; 98US-0103310.
XX PA (REGC ) UNIV CALIFORNIA.
XX PI Masliah E;
XX DR WPI; 2000-303637/26.
XX PT Screening treatments for neurodegenerative disease, such as Alzheimer's
XX PT disease, comprises stimulating oxidative stress and measuring and
XX PT aggregation of non-amyloid beta component /alpha-synuclein in neuronal
XX PT cells
XX PS Example 4; Page 47; 47pp; English.
XX CC The invention concerns methods for screening treatments for, and
XX CC treatment of, neurodegenerative disease. The methods comprise measuring
XX CC aggregation of neurons of NACP/alpha-synuclein and stimulating expression
XX CC of a non-amyloidogenic protein in order to reduce the level of
XX CC aggregation. In the screening method, oxidative stress is stimulated in
XX CC the neuronal cells by introduction of metal ions and hydrogen peroxide.
XX CC The methods are useful for treating and screening treatments for
XX CC neurodegenerative disease consisting of Alzheimer's disease, Parkinson's
XX CC disease and Lewy body disease.
XX SQ Sequence 140 AA;
Query Match 94.4%; Score 663; DB 21; Length 140;
Best Local Similarity 94.3%; Pred. No. 2.3e-56;
Matches 132; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 MDVFMKGLSKAKGEGVAAAEKTKGVAAAGKTKGVLVYVGSKTKGCVHGVATVAEKK 60
DB 1 MDVFMKGLSKAKGEGVAAAEKTKGVAAAGKTKGVLVYVGSKTKGCVHGVATVAEKK 60

QY 61 EQVTNNGGAVVTGVTAVAQKTVEGAGSIAAATGFKVKKDQLGKNEEGAPQEGILEDMPVDP 120
DB 61 EQVTNNGGAVVTGVTAVAQKTVEGAGSIAAATGFKVKKDQMGKGEYPOEGILEDMPVDP 120

```


OY 121 DNEAYEMPSEEGYQDYPEA 140
:|||||
Db 121 GSEAYEMPSEEGYQDYPEA 140

Search completed: January 28, 2003, 14:13:58
Job time : 8.80115 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:19:14 : Search time 1.83556 Seconds
(without alignments)
1539.035 Million cell updates/sec

Title: US-09-904-987-4
Perfect score: 702
Sequence: 1 MDVFMKGLSKAKGEGVVAEAE.....DNEAYEMPSEGYQDYEPFA 140

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 122226 seqs, 20178551 residues
Total number of hits satisfying chosen parameters: 122226

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Watch 100%
Listing first 45 summaries

Database : Published_Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pbp.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pbp.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pbp.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pbp.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pbp.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pbp.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pbp.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pbp.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pbp.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pbp.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pbp.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pbp.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pbp.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pbp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	702	100.0	140	10	US-09-904-987-4
2	702	100.0	140	12	US-10-039-413-1
3	695	99.0	140	12	US-10-039-413-2
4	692	98.6	140	12	US-10-039-413-3
5	684	97.4	140	12	US-10-039-413-4
6	226	32.2	47	10	US-09-864-761-45900
7	151	21.5	28	10	US-09-864-761-46299
8	143	20.4	30	10	US-09-864-761-47813
9	83	11.8	321	10	US-09-967-347-12
10	80.5	11.5	216	8	US-08-450-842-8
11	79.5	11.3	129	10	US-09-815-242-11228
12	79	11.3	599	10	US-09-789-561-120
13	78	11.1	267	10	US-09-852-555-2
14	77	11.0	561	10	US-09-812-079A-2
15	76	10.8	328	10	US-09-765-272-8
16	76	10.8	475	10	US-09-734-569-166
17	76	10.8	1599	9	US-10-092-880-9
18	75.5	10.8	202	10	US-09-815-242-5604
19	75.5	10.8	203	10	US-09-815-242-12176

20	75	10.7	331	9	US-09-573-446-1	Sequence 1, Appli
21	75	10.7	487	10	US-09-745-763-108	Sequence 108, App
22	74.5	10.6	413	9	US-09-738-626-4478	Sequence 4478, Ap
23	73.5	10.5	714	10	US-09-978-242-3	Sequence 3, Appli
24	73	10.4	1781	10	US-09-738-877-3	Sequence 3, Appli
25	72.5	10.3	878	10	US-09-815-242-13957	Sequence 13957, A
26	72	10.3	401	10	US-09-815-242-13828	Sequence 13828, A
27	72	10.3	568	10	US-09-815-242-5140	Sequence 5140, Ap
28	72	10.3	836	9	US-09-738-626-5805	Sequence 5805, Ap
29	72	10.3	1536	9	US-10-092-880-2	Sequence 2, Appli
30	71.5	10.2	709	10	US-09-815-242-10989	Sequence 10989, A
31	71	10.1	811	10	US-09-840-787-45	Sequence 45, Appl
32	71	10.1	1002	9	US-09-988-117-3	Sequence 3, Appli
33	71	10.1	1002	10	US-09-812-471-3	Sequence 3, Appli
34	71	10.1	1002	10	US-09-812-633-3	Sequence 3, Appli
35	71	10.1	1477	9	US-10-092-880-4	Sequence 4, Appli
36	70.5	10.0	149	10	US-09-815-242-13554	Sequence 13554, A
37	70	10.0	338	10	US-09-815-242-13457	Sequence 13457, A
38	70	10.0	338	10	US-09-815-242-13645	Sequence 13645, A
39	70	10.0	486	10	US-09-825-414-7	Sequence 7, Appli
40	70	10.0	978	10	US-09-815-242-5456	Sequence 5456, Ap
41	70	10.0	1001	10	US-09-815-242-12686	Sequence 12686, A
42	70	10.0	1600	9	US-10-092-880-10	Sequence 10, Appl
43	70	10.0	2478	10	US-09-815-242-5816	Sequence 5816, Ap
44	70	10.0	2478	10	US-09-815-242-12967	Sequence 12967, A
45	69.5	9.9	484	10	US-09-815-242-10363	Sequence 10363, A

ALIGNMENTS

RESULT 1
US-09-904-987-4
; Sequence 4, Application US/09904987
; Patent No. US20020037908A1
; GENERAL INFORMATION:
; APPLICANT: No. US20020037908Alactyl, Inc.
; TITLE OF INVENTION: Methods and Compositions for Controlling Pathological and Prep
; TITLE OF INVENTION: Protein Assembly or Aggregation
; FILE REFERENCE: 42108/26146
; CURRENT APPLICATION NUMBER: US/09/904,987
; CURRENT FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 140
; TYPE: PRT
; ORGANISM: homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NCBI ENTREZ / XM_003494
; DATABASE ENTRY DATE: 2001-04-16
; RELEVANT RESIDUES: (1)..(140)
US-09-904-987-4

Query Match	100.0%	Score	702	DB	10	Length	140
Best Local Similarity	100.0%	Pred. No.	3.2e-58				
Matches	140	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
Qy	1	MDVFMKGLSKAKGEGVVAEAEKTKQGVAAAGTKGKGVLYVSGTKGKGVVHGVAATVAEKTK	60				
Db	1	MDVFMKGLSKAKGEGVVAEAEKTKQGVAAAGTKGKGVLYVSGTKGKGVVHGVAATVAEKTK	60				
Qy	61	EQVTNNGAVVTGVTAVAAQKTVEGAGSATAATGFVKKQDLGKNREGAPOEGILEDMPVDP	120				
Db	61	EQVTNNGAVVTGVTAVAAQKTVEGAGSATAATGFVKKQDLGKNREGAPOEGILEDMPVDP	120				
Qy	121	DNEAYEMPSEGYQDYEPFA	140				
Db	121	DNEAYEMPSEGYQDYEPFA	140				
RESULT 2							
US-10-039-413-1							

```
; Sequence 1, Application US/10039413
; Patent No. US20020152480A1
; GENERAL INFORMATION:
; APPLICANT: Biere, Anja Leona
; APPLICANT: Citron, Martin
; TITLE OF INVENTION: ALPHA-SYNUCLEIN SUPER-MUTANTS ACCELERATE
; TITLE OF INVENTION: ALPHA-SYNUCLEIN AGGREGATION
; FILE REFERENCE: A-565
; CURRENT APPLICATION NUMBER: US/10/039,413
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 60/101,862
; PRIOR FILING DATE: 1998-09-25
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 140
; TYPE: PRT
; ORGANISM: ADULT HUMAN BRAIN
US-10-039-413-1
Query Match 100.0%; Score 702; DB 12; Length 140;
Best Local Similarity 100.0%; Pred. No. 3.2e-58;
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDVFMKGLSKAKGVVAAAEKTKQGVAAEAGTKEGVLYVGSKTKEGVHGVATVAEKT 60
Db 1 MDVFMKGLSKAKGVVAAAEKTKQGVAAEAGTKEGVLYVGSKTKEGVHGVATVAEKT 60
QY 61 EQVTNVGGAVVTGTTVAQAQTVGAGSIAAATGFKKDLGKNEEGAPQEGILEDMPVDP 120
Db 61 EQVTNVGGAVVTGTTVAQAQTVGAGSIAAATGFKKDLGKNEEGAPQEGILEDMPVDP 120
QY 121 DNEAYEMPSEEGYQDYEP 140
Db 121 DNEAYEMPSEEGYQDYEP 140

RESULT 3
US-10-039-413-2
; Sequence 2, Application US/10039413
; Patent No. US20020152480A1
; GENERAL INFORMATION:
; APPLICANT: Biere, Anja Leona
; APPLICANT: Citron, Martin
; TITLE OF INVENTION: ALPHA-SYNUCLEIN SUPER-MUTANTS ACCELERATE
; TITLE OF INVENTION: ALPHA-SYNUCLEIN AGGREGATION
; FILE REFERENCE: A-565
; CURRENT APPLICATION NUMBER: US/10/039,413
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 60/101,862
; PRIOR FILING DATE: 1998-09-25
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 140
; TYPE: PRT
; ORGANISM: ADULT HUMAN BRAIN
US-10-039-413-2
Query Match 99.0%; Score 695; DB 12; Length 140;
Best Local Similarity 98.6%; Pred. No. 1.4e-57;
Matches 138; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDVFMKGLSKAKGVVAAAEKTKQGVAAEAGTKEGVLYVGSKTKEGVHGVATVAEKT 60
Db 1 MDVFMKGLSKAKGVVAAAEKTKQGVAAEAGTKEGVLYVGSKTKEGVHGVATVAEKT 60
QY 61 EQVTNVGGAVVTGTTVAQAQTVGAGSIAAATGFKKDLGKNEEGAPQEGILEDMPVDP 120
Db 61 EQVTNVGGAVVTGTTVAQAQTVGAGSIAAATGFKKDLGKNEEGAPQEGILEDMPVDP 120
QY 121 DNEAYEMPSEEGYQDYEP 140
Db 121 DNEAYEMPSEEGYQDYEP 140

RESULT 4
US-10-039-413-3
; Sequence 3, Application US/10039413
; Patent No. US20020152480A1
; GENERAL INFORMATION:
; APPLICANT: Biere, Anja Leona
; APPLICANT: Citron, Martin
; TITLE OF INVENTION: ALPHA-SYNUCLEIN SUPER-MUTANTS ACCELERATE
; TITLE OF INVENTION: ALPHA-SYNUCLEIN AGGREGATION
; FILE REFERENCE: A-565
; CURRENT APPLICATION NUMBER: US/10/039,413
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 60/101,862
; PRIOR FILING DATE: 1998-09-25
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 140
; TYPE: PRT
; ORGANISM: ADULT HUMAN BRAIN
US-10-039-413-3
Query Match 98.6%; Score 692; DB 12; Length 140;
Best Local Similarity 98.6%; Pred. No. 2.7e-57;
Matches 138; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDVFMKGLSKAKGVVAAAEKTKQGVAAEAGTKEGVLYVGSKTKEGVHGVATVAEKT 60
Db 1 MDVFMKGLSKAKGVVAAAEKTKQGVAAEAGTKEGVLYVGSKTKEGVHGVATVAEKT 60
QY 61 EQVTNVGGAVVTGTTVAQAQTVGAGSIAAATGFKKDLGKNEEGAPQEGILEDMPVDP 120
Db 61 EQVTNVGGAVVTGTTVAQAQTVGAGSIAAATGFKKDLGKNEEGAPQEGILEDMPVDP 120
QY 121 DNEAYEMPSEEGYQDYEP 140
Db 121 DNEAYEMPSEEGYQDYEP 140

RESULT 5
US-10-039-413-4
; Sequence 4, Application US/10039413
; Patent No. US20020152480A1
; GENERAL INFORMATION:
; APPLICANT: Biere, Anja Leona
; APPLICANT: Citron, Martin
; TITLE OF INVENTION: ALPHA-SYNUCLEIN SUPER-MUTANTS ACCELERATE
; TITLE OF INVENTION: ALPHA-SYNUCLEIN AGGREGATION
; FILE REFERENCE: A-565
; CURRENT APPLICATION NUMBER: US/10/039,413
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 60/101,862
; PRIOR FILING DATE: 1998-09-25
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 140
; TYPE: PRT
; ORGANISM: ADULT HUMAN BRAIN
US-10-039-413-4
Query Match 97.4%; Score 684; DB 12; Length 140;
Best Local Similarity 97.9%; Pred. No. 1.5e-56;
Matches 137; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MDVFMKGLSKAKGVVAAAEKTKQGVAAEAGTKEGVLYVGSKTKEGVHGVATVAEKT 60
Db 1 MDVFMKGLSKAKGVVAAAEKTKQGVAAEAGTKEGVLYVGSKTKEGVHGVATVAEKT 60
QY 61 EQVTNVGGAVVTGTTVAQAQTVGAGSIAAATGFKKDLGKNEEGAPQEGILEDMPVDP 120
Db 61 EQVTNVGGAVVTGTTVAQAQTVGAGSIAAATGFKKDLGKNEEGAPQEGILEDMPVDP 120
```

Db 61 EQVTNNGGAVTGTAVTAOKTVGAGSIAAATGCFVKKDDGRNEGAPCEGILEDPVDP 120
Qy 121 DNEAYEMPSEEGYQDYEP 140
Db 121 DNEAYEMPSEEGYQDYEP 140

RESULT 6

US-09-864-761-45900
; Sequence 45900, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 45900
; LENGTH: 47
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC015529.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.8
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.3
; OTHER INFORMATION: EST_HUMAN HIT: H70324.1, EVALUE 1.00e-08
; OTHER INFORMATION: SWISSPROT HIT: P37840, EVALUE 9.00e-10
US-09-864-761-45900

Query Match 32.2%; Score 226; DB 10; Length 47;
Best Local Similarity 100.0%; Pred. No. 7.6e-15;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 56 AEKTKQVTNNGGAVTGTAVTAOKTVGAGSIAAATGCFVKKDDQLGK 102
Db 1 AEKTKQVTNNGGAVTGTAVTAOKTVGAGSIAAATGCFVKKDDQLGK 47

RESULT 7

US-09-864-761-46299
; Sequence 46299, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 46299
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC011025.4
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.4
; OTHER INFORMATION: EST_HUMAN HIT: AW793306.1, EVALUE 8.00e-10
; OTHER INFORMATION: SWISSPROT HIT: P37840, EVALUE 6.00e-11
US-09-864-761-46299

;; PRIOR APPLICATION NUMBER: 60/242,578

```

8  LSKAKEGVAAAE-----KTKQVABAAAGKREGVLYVG-SKT-----KEGVV 49
      : ||| :
327 LNDVSEGVTGLIKYGVNGGLIRNVTHGVSNSAKGF-AGTISDLGKTMDSNRHQSEREYIR 385
      : ||| :
50  HCVATVAEKTKEQVTNNVGAVVTGTVATAQKTVAGSAGSIAAATGFVKDQKNGEBA-- 107
      : || : ||| : ||| :
386 YHAATSGEHLVAGTHGLAHGIIGLTSVITSTVEGVKTEGVSGFI--SGLGKLGVTPT 443
      : ||| :
108 -PQEGILE I14
      | | | :
444 KPVAGALD 451
      | | | :

```

```
RESULT 13
US-09-852-555-2
; Sequence 2, Application US/09852555
; Patent No. US20020034751A1
; GENERAL INFORMATION:
; APPLICANT: Yuen, Kwok-Yung
; APPLICANT: Cao, Liang
; TITLE OF INVENTION: Aspergillus Fumigatus Antigenic Protein 1
; FILE REFERENCE: 609920-600015
; CURRENT APPLICATION NUMBER: US/09/852,555
; CURRENT FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: US 60/203,322
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Aspergillus Fumigatus
US-09-852-555-2
Query Match 11.1%; Score 78; DB 10; Length 267;
Best Local Similarity 26.4%; Pred. No. 3;
Matches 24; Conservative 20; Mismatches 41; Indels 6; Gaps 3;
QY 5 MKGLSKAKGVV-AAEKTQGVAAEAGTKGVLVVGSKTKEGVHVA-TVAEKTKEQ 62
Db 76 VQDTUKQVEGVDDLSKKDKFVAANG-----GTVYEDLKAQYTAADSLAKAISAKVPES 131
QY 63 VTNVGGAVVTGVTAVAQKTVGAGSTAAATG 93
Db 132 LSDIAAQLSAGITAAIQKIDAYKDAASSTG 162
RESULT 14
US-09-812-079A-2
; Sequence 2, Application US/09812079A
; Patent No. US20020058038A1
; GENERAL INFORMATION:
; APPLICANT: Nicolette, Charles A.
; TITLE OF INVENTION: THERAPEUTIC ANTI-CYTOMEGALOVIRUS
; FILE REFERENCE: GZ 209500
; CURRENT APPLICATION NUMBER: US/09/812,079A
; CURRENT FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/191,050
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/254,989
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 561
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-812-079A-2
Query Match 11.0%; Score 77; DB 10; Length 561;
Best Local Similarity 26.5%; Pred. No. 9.2;
Matches 26; Conservative 17; Mismatches 45; Indels 10; Gaps 3;
QY 28 EAAGTKEGVLVVGSKTKEGVHVAETKEQVTNNGAVVTGVTAVAQKTVGAGS 87
Db 390 EGAAGDDDDVTSGSDSEELV-----TTERKTPRVTG-GGAMAGASTSAGRKRKSASSA 443
QY 88 TAAATGFKVKDQKNEGAPQEGILEMDPVPDPNEAY 125
Db 444 TACTSGVMTRGLRAESTVAPE-----EDTDEDSNEIH 477
RESULT 15
US-09-765-272-8
; Sequence 8, Application US/09765272
```

```
; Patent No. US20020061545A1
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/765,272
; FILING DATE: 22-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 328 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-765-272-8
Query Match 10.8%; Score 76; DB 10; Length 328;
Best Local Similarity 30.8%; Pred. No. 5.9;
Matches 36; Conservative 12; Mismatches 55; Indels 14; Gaps 5;
QY 16 VAAAEKTKQGVAAEAGTKGVLVVGSKTKEGVHVAETKEQVTNNG--GAVVTG 73
Db 132 VAAAKTTTKQGVFGVGGIESEVI---SREFAGFKAGVASVDFSIKQVDYAGSFGDAAKG 188
QY 74 VTAVAQKTVGAG---STAAATG---FVKKDLGKNEGAPQEGILEMDPVPDPNEA 124
Db 189 KTIATAAQYAAGADIVYQVAGTGAGVFAEAKSL---NESRPENKVVWVIGVDRDQEA 242
Search completed: January 28, 2003, 14:40:08
Job time : 2.83556 secs
```


GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:10:57 ; Search time 2.90631 Seconds
(without alignments)
1417.333 Million cell updates/sec

Title: US-09-904-987-4

Perfect score: 702

Sequence: 1 MDVFMKGLSKAKEGVVAAAE.....DNEAYEMPSEGYQDYEPFA 140

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*

- 1: /cgn2.6/ptodata/1/1aa/5A_COMB.pep:*
- 2: /cgn2.6/ptodata/1/1aa/5B_COMB.pep:*
- 3: /cgn2.6/ptodata/1/1aa/6A_COMB.pep:*
- 4: /cgn2.6/ptodata/1/1aa/6B_COMB.pep:*
- 5: /cgn2.6/ptodata/1/1aa/PCTUS_COMB.pep:*
- 6: /cgn2.6/ptodata/1/1aa/backfiles1.pep:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	702	100.0	140	US-09-405-035-1	Sequence 1, Appli
2	702	100.0	140	PCT-US94-09789-2	Sequence 2, Appli
3	695	99.0	140	US-09-405-035-2	Sequence 2, Appli
4	692	98.6	140	US-09-405-035-3	Sequence 3, Appli
5	694	97.4	140	US-09-405-035-4	Sequence 4, Appli
6	307.5	43.8	127	US-08-705-771-12	Sequence 12, Appli
7	166	23.6	35	PCT-US94-09789-3	Sequence 3, Appli
8	95	13.5	20	PCT-US94-09789-4	Sequence 4, Appli
9	93	13.2	216	US-09-125-619-46	Sequence 46, Appli
10	88	12.5	214	US-09-125-619-34	Sequence 34, Appli
11	87	12.4	190	US-09-125-619-15	Sequence 15, Appli
12	87	12.4	212	US-09-125-619-31	Sequence 31, Appli
13	87	12.4	212	US-09-125-619-43	Sequence 43, Appli
14	87	12.4	356	US-09-125-619-2	Sequence 2, Appli
15	87	12.4	356	US-09-125-619-13	Sequence 13, Appli
16	85.5	12.2	189	US-09-125-619-26	Sequence 26, Appli
17	83	11.8	321	US-09-171-5178-12	Sequence 12, Appli
18	82.5	11.8	187	US-09-125-619-23	Sequence 23, Appli
19	80.5	11.5	190	US-09-125-619-19	Sequence 19, Appli
20	80.5	11.5	216	US-08-451-947-8	Sequence 8, Appli
21	80.5	11.5	216	US-08-424-826A-8	Sequence 8, Appli
22	80.5	11.5	216	US-08-928-694-8	Sequence 8, Appli
23	80.5	11.5	216	PCT-US91-06950-8	Sequence 8, Appli
24	79.5	11.3	211	US-09-125-619-40	Sequence 40, Appli
25	79.5	11.3	212	US-09-125-619-41	Sequence 41, Appli
26	79	11.3	194	US-09-125-619-28	Sequence 28, Appli
27	78.5	11.2	281	US-08-405-175A-9	Sequence 9, Appli

28	78	11.1	212	4	US-09-125-619-32	Sequence 32, Appli
29	77.5	11.0	179	4	US-09-125-619-18	Sequence 18, Appli
30	77.5	11.0	209	4	US-09-125-619-42	Sequence 42, Appli
31	77.5	11.0	729	4	US-08-887-534A-30	Sequence 30, Appli
32	77.5	11.0	1004	4	US-09-268-347-30	Sequence 30, Appli
33	77	11.0	579	4	US-09-171-699-8	Sequence 8, Appli
34	76.5	10.9	215	4	US-09-125-619-36	Sequence 36, Appli
35	76.5	10.9	892	4	US-09-336-447A-5	Sequence 5, Appli
36	76.5	10.9	1477	4	US-09-206-942-71	Sequence 71, Appli
37	76	10.8	158	4	US-09-125-619-21	Sequence 21, Appli
38	76	10.8	195	4	US-09-125-619-17	Sequence 17, Appli
39	76	10.8	209	4	US-09-125-619-33	Sequence 33, Appli
40	76	10.8	211	4	US-09-125-619-37	Sequence 37, Appli
41	76	10.8	213	4	US-09-125-619-39	Sequence 39, Appli
42	76	10.8	328	4	US-08-961-083-8	Sequence 8, Appli
43	76	10.8	579	4	US-09-171-699-6	Sequence 6, Appli
44	76	10.8	1338	2	US-08-728-470-9	Sequence 9, Appli
45	76	10.8	1338	4	US-08-719-641-9	Sequence 9, Appli

ALIGNMENTS

RESULT 1
US-09-405-035-1
; Sequence 1, Application US/09405035
; Patent No. 6184351
; GENERAL INFORMATION:
; APPLICANT: Biere, Anja Leona
; APPLICANT: Citron, Martin
; TITLE OF INVENTION: ALPHA-SYNUCLEIN SUPER-MUTANTS ACCELERATE
; TITLE OF INVENTION: ALPHA-SYNUCLEIN AGGREGATION
; FILE REFERENCE: A-565
; CURRENT APPLICATION NUMBER: US/09/405.035
; CURRENT FILING DATE: 1999-09-24
; EARLIER APPLICATION NUMBER: 60/101,862
; EARLIER FILING DATE: 1998-09-25
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 1
; LENGTH: 140
; TYPE: PRT
; ORGANISM: ADULT HUMAN BRAIN
US-09-405-035-1

Query Match 100.0%; Score 702; DB 4; Length 140;
Best Local Similarity 100.0%; Pred. No. 3.9e-68;
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MDVFMKGLSKAKEGVVAAAEKTKQGVAAAGTKEGVLYVGSKTKEGVVHGVATVAEKT 60
Db 1 MDVFMKGLSKAKEGVVAAAEKTKQGVAAAGTKEGVLYVGSKTKEGVVHGVATVAEKT 60
Qy 61 EQVTNVGAVTGTAVTAQKTVEGAGSTAAATGFVKDQDKNEEGAPQEGILEMPVDP 120
Db 61 EQVTNVGAVTGTAVTAQKTVEGAGSTAAATGFVKDQDKNEEGAPQEGILEMPVDP 120
Qy 121 DNEAYEMPSEGYQDYEPFA 140
Db 121 DNEAYEMPSEGYQDYEPFA 140

RESULT 2
PCT-US94-09789-2
; Sequence 2, Application PC/TUS9409789
; GENERAL INFORMATION:
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: NOVEL COMPONENT OF AMYLOID IN
; TITLE OF INVENTION: ALZHEIMER'S DISEASE AND METHODS FOR USE OF SAME
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 1880 Century Park East - Suite 500


```

; LENGTH: 140
; TYPE: PRT
; ORGANISM: AN
US-09-405-035-4

```

Query Match	97.4%;	Score 684;	DB 4;	Length 140;
Best Local Similarity	97.9%;	Pred. No. 3.3e-66;		
Matches 137;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps
QY	1	MDVFMKGLSKAKGVAAAEKTKQGVAAEAGTKKEGLVYVSGSKTEGVVHGVAIVAEKTK 60		
Db	1	MDVFMKGLSKAKGVAAAEKTKQGVAAEAGTKKEGLVYVSGSKTEGVVHGVAIVAEKTK 60		
QY	61	EQVTVNGGAVVTVTAQAQTVESGAGSIAAATGFFVKDQQLGKNBEGAPQSGILEDMPVDP 120		
Db	61	EQVTVNGGAVVTVTAQAQTVESGAGSIAAATGFFVKDQQLGKNBEGAPQSGILEDMPVDP 120		
QY	121	DNEAYEMPSEGGYQDYDEPEA 140		
Db	121	DNEAYEMPSEGGYQDYDEPEA 140		

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Query Match	43.8%	Score 307.5	DB 3	Length 127
Best Local Similarity	61.8%	Pred. No. 8.5e-26		
Matches 68	Conservative 10	Mismatches 29	Indels 3	Gaps
1 MDVFMGLSKAKEGVVAAAEKTKQGVAAAGTKTEGLVYGSKTKEGVHGATVAEATK 60				
1 MDVFMGLSKAKEGVVAAAEKTKQGVAAAGTKTEGLVYGSKTKEGVHGATVAEATK 60				
61 EQVNTNNGGAVVTCGTAVAQKTEVGAGSIAAATGTFVKDQLQKNEGAPOE 110				

Db 61 EQANAVSKAVSSVNTVATKTVEEAENIAVTSGVVRKEDL---RPSAPQ 107

```

RESULT 7
PCT-US94-09789-3
; Sequence 3, Application PC/TUS9409789
; GENERAL INFORMATION:
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: NOVEL COMPONENT OF AMYLOID IN
; TITLE OF INVENTION: ALZHEIMER'S DISEASE AND METHODS FOR
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 1880 Century Park East - Suite 500
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/09789
; FILING DATE: 29-AUG-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Howells, Stacy L.
; REGISTRATION NUMBER: 34,842
; REFERENCE/DOCKET NUMBER: FD-3520
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 455-5100
; TELEFAX: (619) 455-5110
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; CLONE: NAC
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..35
; PCT-US94-09789-3

```

```

Query Match      23.6%; Score 166; DB 5; Length 35;
Best Local Similarity 100.0%; Pred. No. 2.4e-11;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OQY 61 EQVTNNVGAVVTGVTAVAQKTVEGAGSIAAATGFV 95
      |||||
Db 1 EQVTNNVGAVVTGVTAVAQKTVEGAGSIAAATGFV 35
      |||||

```

RESULT 8
PCT-US94-09789-4
; Sequence 4, Application PC/TUS9409789
; GENERAL INFORMATION:
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: NOVEL COMPONENT OF AMYLOID IN
; TITLE OF INVENTION: ALZHEIMER'S DISEASE AND METHODS FOR USE OF SAME
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 1880 Century Park East - Suite 500
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90067

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/09789
FILING DATE: 29-AUG-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Howells, Stacy L.
REGISTRATION NUMBER: 34,842
REFERENCE/DOCKET NUMBER: FD-3520
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 455-5100
TELEFAX: (619) 455-5110
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
CLONE: Peptide X
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..20
PCT-US94-09789-4

Query Match 13.5%; Score 95; DB 5; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.00047;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 61 EQVTNVGAVVTGVTAAQK 80
Db 1 EQVTNVGAVVTGVTAAQK 20

RESULT 9
US-09-125-619-46
; Sequence 46, Application US/09125619
; Patent No. 6437116
; GENERAL INFORMATION:
; APPLICANT: NORRIS, STEVEN J.
; APPLICANT: JING-REN, ZHANG
; APPLICANT: HARDHAM, JOHN M.
; APPLICANT: HOWELL, JERRILYN K.
; APPLICANT: BARBOUR, ALAN G.
; TITLE OF INVENTION: VMP-LIKE SEQUENCES OF PATHOGENIC BORRELIA
; FILE REFERENCE: UTSH:234
; CURRENT APPLICATION NUMBER: US/09/125,619
; CURRENT FILING DATE: 1999-01-27
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 46
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-09-125-619-46

Query Match 13.2%; Score 93; DB 4; Length 216;
Best Local Similarity 29.7%; Pred. No. 0.02;
Matches 38; Conservative 19; Mismatches 45; Indels 26; Gaps 6;

Qy 5 MKGLSKAKGVVAAEKTQGVAAEAKTKK 60
Db 63 VKGIAGIKIIVEAAGSEKLVAAAKTNSKAGKLF---GRVDGAAGHDSAAASKAA 119
Qy 61 EQVTNVGG-----AVVTGTAVA-----OKTVEGAGSIAAATGFVKKDLGKNEEGAP 108
Db 120 GAVSVSGEQLSAIVTAAAGAAAEQDGKPKADATNPAAAAIG-----KGNEENGAE 173

Qy 109 -QEGILED 115
Db 174 FGDGMKKD 181

RESULT 10
US-09-125-619-34
; Sequence 34, Application US/09125619
; Patent No. 6437116
; GENERAL INFORMATION:
; APPLICANT: NORRIS, STEVEN J.
; APPLICANT: JING-REN, ZHANG
; APPLICANT: HARDHAM, JOHN M.
; APPLICANT: HOWELL, JERRILYN K.
; APPLICANT: BARBOUR, ALAN G.
; APPLICANT: WEINSTOCK, GEORGE M.
; TITLE OF INVENTION: VMP-LIKE SEQUENCES OF PATHOGENIC BORRELIA
; FILE REFERENCE: UTSH:234
; CURRENT APPLICATION NUMBER: US/09/125,619
; CURRENT FILING DATE: 1999-01-27
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 34
; LENGTH: 214
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-09-125-619-34

Query Match 12.5%; Score 88; DB 4; Length 214;
Best Local Similarity 32.0%; Pred. No. 0.067;
Matches 39; Conservative 14; Mismatches 49; Indels 20; Gaps 6;

Qy 7 GLSKA-KEGVAAAEKTKQGVAAEAKTKK 65
Db 65 GIAGIKIIVEAAGSEKLVAAAKENKAGKLF 123
Qy 66 VGG-----AVVTGTAVA-----OKTVEGAGSIAAATGFVKKDLGKNEEGAP--QEGIL 113
Db 124 VSGEQLSAIVTAAAGAAAEQDGKPKAEAKNPAAAAIG-----KGNEENGAEFNKQGMK 177

Qy 114 ED 115
Db 178 KD 179

RESULT 11
US-09-125-619-15
; Sequence 15, Application US/09125619
; Patent No. 6437116
; GENERAL INFORMATION:
; APPLICANT: NORRIS, STEVEN J.
; APPLICANT: JING-REN, ZHANG
; APPLICANT: HARDHAM, JOHN M.
; APPLICANT: HOWELL, JERRILYN K.
; APPLICANT: BARBOUR, ALAN G.
; APPLICANT: WEINSTOCK, GEORGE M.
; TITLE OF INVENTION: VMP-LIKE SEQUENCES OF PATHOGENIC BORRELIA
; FILE REFERENCE: UTSH:234
; CURRENT APPLICATION NUMBER: US/09/125,619
; CURRENT FILING DATE: 1999-01-27
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-09-125-619-15

Query Match 12.4%; Score 87; DB 4; Length 190;
Best Local Similarity 33.3%; Pred. No. 0.073;
Matches 38; Conservative 14; Mismatches 44; Indels 18; Gaps 7;

5	MKGLSKAKGCVAAA	---	EKTKQGVAEAGKTKGGVLYVSGKTKG	VGVHGVATVAETKE	61
49	VKGIAGKIBIVEAAGSGSEKLK	-	AVAAAKGNNNGAGKLGCKA	-	GAAAHGDSAAKRAAG 106
62	QVTNWGG	----	AVVTGVTVAQA	----	KTVGAGSTAATGTFVKKD --- QLGKNE 104
107	AVSAVSGPOTII	SAIVTAADAAEODGCKP	PEERAKNP	IAAAG -	DKDGGAFGGODE 158

```

RESULT 12
US-09-125-619-31
; Sequence 31, Application US/09125619
; Patent No. 6437116
; GENERAL INFORMATION:
; APPLICANT: NORRIS, STEVEN J.
; APPLICANT: JING-REN, ZHANG
; APPLICANT: HARDHAM, JOHN M.
; APPLICANT: HOWELL, JERRILYN K.
; APPLICANT: BARBOUR, ALAN G.
; APPLICANT: WEINSTOCK, GEORGE M.
; TITLE OF INVENTION: VMP-LIKE SEQUENCES OF PATHOGENIC BORRELIA
; FILE REFERENCE: UTSH:234
; CURRENT APPLICATION NUMBER: US/09/125,619
; CURRENT FILING DATE: 1999-01-27
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 31
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-09-125-619-31

```

```
Query Watch      12.48; Score 87; DB 4; Length 212;  
Best Local Similarity 33.38; Pred. No. 0.085;  
Matches 38; Conservative 14; Mismatches 44; Indels 18; Gaps 7;  
  
QY    5   MKGLSKAKGGVVAAA---EKKTKQGVAAGAAKTEGVLYVGSCTKECVGHGATVAERTKE 61  
       |||::||::|||::|||::|||::|||::|||::|||:  
Db     65  VKIANGIKRIVEAAGSGSEKLK-AVAAAKGENNKGAKLFGKA-GAAAHGDESAASKRAAG 122  
                               |  
QY     62  QVTNVGG-----AVVTGTVAQAQ---RTVSEGASIAAANTGFVKKD---QLGRNE 104  
                               |  
Db     123 AVSAVSGEOILSAIVTAADAAEDCGKPPEAKNPAAAAIG--DKDGAAGFEQODE 174
```

```

RESULT 13
US-09-125-619-43
; Sequence 43, Application US/09125619
; Patent No. 6437116
; GENERAL INFORMATION:
; APPLICANT: NORRIS, STEVEN J.
; APPLICANT: JING-REN, ZHANG
; APPLICANT: HARDHAM, JOHN M.
; APPLICANT: HOWELL, JERRILYN K.
; APPLICANT: BARBOUR, ALAN G.
; APPLICANT: WEINSTOCK, GEORGE M.
; TITLE OF INVENTION: VMP-LIKE SEQUENCES OF PATHOGENIC BORRELIA
; FILE REFERENCE: UTSH:234
; CURRENT APPLICATION NUMBER: US/09/125.619
; CURRENT FILING DATE: 1999-01-27
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 43
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-09-125-619-43

```

Query Match 12.4%; Score 87; DB 4; Length 212;
Best Local Similarity 33.3%; Pred. No. 0.085;
Matches 36; Conservative 14; Mismatches 44; Indels 18; Gaps 7;
QY 5 MGLGSAKAGGWAAA---EKTQGVAAEAGKTKGVLVYSGTKGCVVHGVATVAEKTKE 61

Db 65 VKGIAGIKETVEAAGSEKUL-AVAAKAGNNKGAKLFGKA-GAAAHGDSAAKAG 122
QY 62 QVTNNGG-----AVVTGTVTAQAQ---KTVEGAGSIAAATGFVKRD---QLCKNE 104
Db 123 AVSAYSGEQLSALVTAADAFAEQDGKPEEAKNPATAAIG---DKDGAEFGQDE 174

```

RESULT 14
US-09-125-619-2
; Sequence 2, Application US/09125619
; Patent No. 6437116
; GENERAL INFORMATION:
; APPLICANT: NOREIS, STEVEN J.
; APPLICANT: JING-REN, ZHANG
; APPLICANT: HARDHAM, JOHN M.
; APPLICANT: HOWELL, JERRILYN K.
; APPLICANT: BARBOUR, ALAN G.
; APPLICANT: WEINSTOCK, GEORGE M.
; TITLE OF INVENTION: VMP-LIKE SEQUENCES OF PATHOGENIC BORRELIA
; FILE REFERENCE: UTSH:234
; CURRENT APPLICATION NUMBER: US/09/125,619
; CURRENT FILING DATE: 1999-01-27
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 356
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
; US-09-125-619-2

```

[illegible]

RESULT 15
US-09-125-619-13
; Sequence 13, Application US/09125619
; Patent No. 6437116
; GENERAL INFORMATION:
; APPLICANT: NORRIS, STEVEN J.
; APPLICANT: JING-REN, ZHANG
; APPLICANT: HARDHAM, JOHN M.
; APPLICANT: HOWELL, JERRILYN K.
; APPLICANT: BARBOUR, ALAN G.
; APPLICANT: WEINSTOCK, GEORGE M.
; TITLE OF INVENTION: VMP-LIKE SEQUENCES OF PATHOGENIC BORRELIA
; FILE REFERENCE: UTSH:234
; CURRENT APPLICATION NUMBER: US/09/125,619
; CURRENT FILING DATE: 1999-01-27
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 356
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-09-125-619-13

Query Match 12.4%; Score 87; DB 4; Length 356;
Best Local Similarity 33.3%; pred. No. 0.17;
Matches 38; Conservative 14; Mismatches 44; Indels 18; Gaps 7;

QY 5 MKGLSKAKEGVYAAA---EKTQGVAAAGKTKGVLVGSKTKEGVVHGVAATKE 61
.....| | | | | | | | | |

Db 164 VKGIAGIKIETVEAAGGSEKLK-AVAAAAGENNKGAGKLFKA-GAAAHGDSEAAKKAAG 221
Qy 62 QVTNWGG-----AVVTGVTVAQ---KTVEGAGSIAAATGFVKKD---OLGKNE 104
Db 222 AVSAVSGEQILSAIVTAADAAEQDGKKPEEAKNPAAAAIG--DKOGGAFFGODE 273

Search completed: January 28, 2003, 14:21:22
Job time : 3.90631 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:08:34 : Search time 19.8765 Seconds
(without alignments)
3666.126 Million cell updates/sec

Title: US-09-904-987-5

Perfect score: 3965

Sequence: 1 MAEPKQEFVEMDHAGTYGL.....SPQLATLADVSASLAKQGL 758

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR-73.*

1: p1r1.*

2: p1r2.*

3: p1r3.*

4: p1r4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2780.5	70.1	733	2 A45301	microtubule-associ
2	2552	64.4	686	2 A38235	microtubule-associ
3	2116.5	53.4	441	1 QRHU71	microtubule-associ
4	1845	46.5	432	2 JS0306	microtubule-associ
5	1845	46.5	448	1 ORBOT1	microtubule-associ
6	1770	44.6	402	1 ORBOT2	microtubule-associ
7	1552	39.1	374	2 S46264	microtubule-associ
8	1415.5	35.7	316	1 ORHU72	microtubule-associ
9	1354.5	34.2	341	2 B28820	microtubule-associ
10	1350.5	34.1	364	2 A28820	microtubule-associ
11	812.5	20.5	1824	1 ORHU71	microtubule-associ
12	807	20.4	1828	2 A40115	microtubule-associ
13	796.5	20.1	1830	2 A37981	microtubule-associ
14	795	20.1	1825	2 S13507	microtubule-associ
15	752	19.0	472	2 I67793	microtubule-associ
16	715	18.0	323	2 A55983	microtubule-associ
17	624	15.7	1152	2 A33183	microtubule-associ
18	621.5	15.7	1125	2 B41206	microtubule-associ
19	576	14.5	381	2 S51375	microtubule-associ
20	545	13.7	1072	2 A37127	microtubule-associ
21	507.5	12.8	982	2 A53253	microtubule-associ
22	482.5	12.2	1224	2 T14007	microtubule-associ
23	481	12.1	125	2 I52650	microtubule-associ
24	341	8.6	66	2 S26663	microtubule-associ
25	309	7.8	2187	2 T30826	nascent polypeptid
26	294	7.4	1464	2 S59856	collagen alpha 1(I
27	290	7.3	1049	1 CGB07S	collagen alpha 1(I
28	285.5	7.2	990	2 I51618	nucleolar phosphop
29	285.5	7.2	1806	1 CGHU1E	collagen alpha 1(X

30 285 7.2 971 2 T19431 hypothetical prote
31 285 7.2 1184 2 G01763 atrophin-1 - human
32 283.5 7.2 779 1 CGB01S collagen alpha 1(I
33 283.5 7.2 1466 1 CGHU7L collagen alpha 1(I
34 283 7.1 1838 1 CGHU1V collagen alpha 1(V
35 282 7.1 1151 2 T18535 high molecular mas
36 281.5 7.1 886 2 I50894 collagen alpha 1(I
37 279.5 7.0 1691 1 S22917 collagen alpha 5(I
38 278 7.0 242 2 T16349 hypothetical prote
39 277.5 7.0 1763 2 S16366 hypothetical prote
40 277.5 7.0 2715 2 T13049 eyelid - fruit fly
41 276 7.0 880 2 D89756 protein T23E7.2b l
42 276 7.0 1274 2 T16251 hypothetical prote
43 276 7.0 1492 2 A40333 collagen alpha 1(I
44 275 6.9 1024 2 S18251 collagen alpha 1(X
45 275 6.9 1549 2 I48103 type VII collagen

ALIGNMENTS

RESULT 1

A45301

microtubule-associated protein tau - mouse

N:Alternate names: microtubule binding protein tau

C:Species: Mus musculus (house mouse)

C>Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 13-Aug-1999

C:Accession: A45301; S31658

R:Couchie, D.; Mavilia, C.; Georgieff, I.S.; Liem, R.K.; Shelanski, M.L.; Nunez, J.

A:Title: Primary structure of high molecular weight tau present in the peripheral ner

A:Reference number: A45301; MUID:92262443; PMID:1374898

A:Accession: A45301

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-733 <COU>

A>Note: this sequence is inconsistent with the nucleotide translation

A>Note: sequence extracted from NCBI backbone (NCBI:102045, NCBI:102046)

R:Kenner, L.; Forstner, M.; Hutter, H.; Hoefler, G.; Kurzbauer, R.; Zatloukal, K.; Kr

A:Description: First observation of mRNA for a tau-protein from murine liver and kidn

A:Reference number: S31658

A:Accession: S31658

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-529-651 <KEN>

A:Cross-references: EMBL:Z12133; NID:g54263; PIDN:CAA78121.1; PID:g388534

C:Superfamily: microtubule-associated protein tau; MAP2/tau repeat homology

C:Keywords: microtubule binding; tandem repeat

F:544-574/Domain: MAP2/tau repeat homology <MT1>

F:575-605/Domain: MAP2/tau repeat homology <MT2>

F:606-636/Domain: MAP2/tau repeat homology <MT3>

F:637-668/Domain: MAP2/tau repeat homology <MT4>

Query Match

70.1%; Score 2780.5; DB 2; Length 733;

Best Local Similarity 74.2%; Pred. No. 7.3e-120;

Matches 569; Conservative 44; Mismatches 111; Indels 43; Gaps 11;

QY 1 MAEPKQEFVEMDHAGTYGLGDRKDGQYTMHODQEGDTDAGLKESPLQTPTTDEGSEEPG 60

Db 1 MANPROEFTMEDHA-----GDTLLDQEGMDHGLKESPPQPPADGAEPEG 49

QY 61 SETSDAKSTTAEVDVAPLVDEGAPKQAAAPHTPEIGTTAEAGIGTTPPSLEDEAG 120

Db 50 SETSDAKSTTAEVDVAPLVDERAPDKQAAAPHTPEIGTTAEAGIGTTPNQEDQAAG 109

QY 121 HVTPQEPGKVGVEGFLREPGLSHLHSLMGMPGAPLLPEGPREATROPSCGTPEDTEG 180

Db 110 HVTPQ----GR--REGQADPLGTSDWTRQQVSSMSGAPLLPQGLREATCPQSGTRPEDIEK 163

QY 181 GRHAPELLKQLGLDHLHQEGFPLKAGKRPKSGKEVEDRDVDDESSQDSSPPSKASPA 240

Db 164 SHPASELLR-----RGPPQKEGWDRLGSEEEVDEDLTVDESS-QDSPPSQASLT 213

QY 241 QDGRPPQ-----TAAREATSIIPGFAEAGIPLVDLSKSVSTEIPASEDPGSPVG--RAKQ 295
 DB 214 PGRAAPQAGSGVCGTASVPLPTEGVSPLPADFEFSKVSATQASQEGGCTGPMSEGH 273
 QY 296 DAPLEFTFHVETTPNVQKEQAHSEHLGRAAPGAPGEGPEA--RGPSLGEDTKEADLPE 353
 DB 274 EAAPLEFTFHVETKASTPKBEQ-----DLEGATVVGVPGEQKAQTPGVSAGKGTKEASLQE 328
 QY 354 PSEKQPAAPRCKPVSRYPQLKARVMSKDKGTGSDDKKAKTSTSSAKTLKNRCLSPK 413
 DB 329 PPGKQPAAGLPGRPVSRVPLKARVA--SKDRGTNDEKKAKTSTPSCAKAPSHRPLCLSP 386
 QY 414 LPTPGSSDPLIOPSSPAVCPSPSPKRVSSVTSRTGSSGAKEMKLGADGKT--KIATP 471
 DB 387 RPTLGSSDPLIKPSSPAVSPATSPKRVSSVTPRNGSPGKQMKLGADGKTGAKIATP 446
 QY 472 RGAAPGQKQANATRIAPAKTPPTSSGEPKSGDRSGYSSPGSPGTPGSRSTRPS 531
 DB 447 RGAASPAQKGTSNATRIAPAKTPPTSPKTPPGSGEPKSGDRSGYSSPGSPGTPGSRSTRPS 506
 QY 532 LPTPTPREPKKVAVVRTPPKSSAKSRLQATAPVMPDLKNVSKIGSTENLKHOPGGK 591
 DB 507 LPTPTPREPKKVAVVRTPPKSSAKSRLQATAPVMPDLKNVSKIGSTENLKHOPGGK 566
 QY 592 VQIINKKLDLSNVQSKGSKDKNIKHVPGGGSVQIYVYKPVVDSKSVTSKGSGLNIHHPGG 651
 DB 567 VQIINKKLDLSNVQSKGSKDKNIKHVPGGGSVQIYVYKPVVDSKSVTSKGSGLNIHHPGG 626
 QY 652 GQVEVSEKLDKDRVQSKIGSLDNIHVPVGGNKKIETHKLTFFRENAAKTDHGAIEVY 711
 DB 627 GQVEVSEKLDKDRVQSKIGSLDNIHVPVGGNKKIETHKLTFFRENAAKTDHGAIEVY 686
 QY 712 KSPVSGDTSPRHLSNVSTGSDIMVDSPLATLADSVASLAKOGL 758
 DB 687 KSPVSGDTSPRHLSNVSTGSDIMVDSPLATLADSVASLAKOGL 733

RESULT 2

A38235
 microtubule-associated protein, 110K tau - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 13-Aug-1999
 C:Accession: A38235
 R:Goedert, M.; Spillantini, M.G.; Crowther, R.A.
 Proc. Natl. Acad. Sci. U.S.A. 89, 1983-1987, 1992
 A:Title: Cloning of a big tau microtubule-associated protein characteristic of the perip
 A:Reference number: A38235; MUID:92179305; PMID:1542696
 A:Accession: A38235
 A:Molecule type: mRNA
 A:Residues: 1-686 <GOE>
 A:Cross-references: GB:M84156; NID:g207157; PIDN:AAA42204.1; PID:g207158
 A:Note: sequence extracted from NCBI backbone (NCBIN:87358, NCBI:87359)
 C:Superfamily: microtubule-associated protein tau; MAP2/tau repeat homology
 C:Keywords: alternative splicing; microtubule binding; tandem repeat
 F:497-527/Domain: MAP2/tau repeat homology <MT1>
 F:528-558/Domain: MAP2/tau repeat homology <MT2>
 F:559-589/Domain: MAP2/tau repeat homology <MT3>
 F:590-621/Domain: MAP2/tau repeat homology <MT4>

Query Watch 64.4%; Score 2552; DB 2: Length 686;
 Best Local Similarity 68.3%; Pred. No. 1.8e-109;
 Matches 531; Conservative 40; Mismatches 95; Indels 112; Gaps 11;

QY 1 MAPPEFVEFVMEHDAGTYGLGRDKQGGYTMHQDQEGDTPDAGLKESPLQTPTEDCSEEPG 60
 DB 1 MAEPREFFTMEDQA-----GDTYMLQDQEGDMDHGLKESPPQPPADDCSEEPG 49
 QY 61 SETSDAKSTPTAEDVTAPLVDGAPGKQAAQAPHTPEGTTAAEAGIGDTPSLDEAAG 120
 DB 50 SETSDAKSTPTAEDVTAPLVDGAPGKQAAQAPHTPEGTTAAEAGIGDTPSLDEAAG 109
 QY 121 HVTQEPESKVVQEGFLREFG-----PPGLS---HQLMSGMPGAPLLPEGPREATRPQ 170

DB 110 HVTQEPQKVEIFSQSLIVEPCRRREGQAPDSCISDWTTHQVQPSMSGAPLPQGLREATHQP 169
 QY 171 SGTGPEDTEGRHAPELLKHQLLDLHQEGPLKAGGKPERPGSKKEVEDRDVDDESSPQ 230
 DB 170 LGTRPEDVERSHPASEL-----LWQESPQ--KEAWGKDRLGSEEEVEDDITMDDESS-Q 219
 QY 231 DSPPSKASPAQDGRPPQT-----AAREATSIIPGFAEAGIPLVDLFSKSVSTEIPASE 283
 DB 220 ESPPSQASLAPGTATPQARSVASGVSGETTSIPGFAEAGIPLPADFFSKVSAEQASP 279
 QY 284 PDGSPVGRG--KGQDAPLEFTFHVETTPNVQKEQAHSEHLGRAAPGAPGEGPEARGPSL 342
 DB 280 PEGGCTGPSBEGHEAEPFTFHEI-----KASAPKEQDLEGATVVGAPAEQKARGPSV 334
 QY 343 GEDTKEADLPEPSKQPAAPRCKPVSRYPQLKARVMSKDKGTGSDDKKAKTSTSSAK 402
 DB 335 GKGTKASLLEPTDKQPAAGLPGRPVSRVPLKARVAGVSKDRGTGNDKKA----- 385
 QY 403 TLKNRPCLSPKLPPTPGSSDPLIOPSSPAVCPSPSPKRVSSVTSRTGSSGAKEMKLGKA 462
 DB 386 -----KGA 388
 QY 463 DGKT--KIATPRGAAPCGQKQANATRIAPAKTPPTSSGEPKSGDRSGYSSPGSP 520
 DB 389 DGKTGAKIATPRGAATPGQKTSNATRIAPAKTPPTSPKTPPGSGEPKSGDRSGYSSPGSP 448
 QY 521 GTPGSRSTRPSLPTPTPREPKKVAVVRTPPKSSAKSRLQATAPVMPDLKNVSKIGST 580
 DB 449 GTPGSRSTRPSLPTPTPREPKKVAVVRTPPKSSAKSRLQATAPVMPDLKNVSKIGST 508
 QY 581 ENLKHOPGGGKQVQIINKKLDLSNVQSKGSKDKNIKHVPGGGSVQIYVYKPVVDSKSVTSKCG 640
 DB 509 ENLKHOPGGGKQVQIINKKLDLSNVQSKGSKDKNIKHVPGGGSVQIYVYKPVVDSKSVTSKCG 568
 QY 641 SLGNIHHPGGGQVEVSEKLDKDRVQSKIGSLDNIHVPVGGNKKIETHKLTFFRENAK 700
 DB 569 SLGNIHHPGGGQVEVSEKLDKDRVQSKIGSLDNIHVPVGGNKKIETHKLTFFRENAK 628
 QY 701 AKTDHGAIEVYKSPVSGDTSPRHLSNVSTGSDIMVDSPLATLADSVASLAKOGL 758
 DB 629 AKTDHGAIEVYKSPVSGDTSPRHLSNVSTGSDIMVDSPLATLADSVASLAKOGL 686

RESULT 3

QRHUT1
 microtubule-associated protein tau, long splice form - human
 N:Alternate names: microtubule-binding protein tau; neurofibrillary tangle protein pa
 N:Contains: microtubule-associated protein tau type II; microtubule-associated protei
 C:Species: Homo sapiens (man)
 C:Date: 30-Jun-1990 #sequence_revision 03-May-1996 #text_change 22-Jun-1999
 C:Accession: J03070; A30217; JN0009; S03796; S26665; S26666; S26667; S17302; A3444;
 R:Goedert, M.; Spillantini, M.G.; Jakes, R.; Rutherford, D.; Crowther, R.A.
 Neuron 3, 519-526, 1989
 A:Title: Multiple isoforms of human microtubule-associated protein tau: sequences and
 A:Reference number: J03070; MUID:90380393; PMID:2484340
 A:Accession: J03070
 A:Molecule type: mRNA
 A:Residues: 1-441 <GOE>
 A:Note: six isoforms are found; the clone httau40 sequence is shown. Residues 45-73, 7
 the clone httau24 sequence lacks inserts 1 and 2; the clone httau37 sequence lacks ins
 R:Goedert, M.; Wischik, C.M.; Crowther, R.A.; Walker, J.E.; Klug, A.
 Proc. Natl. Acad. Sci. U.S.A. 85, 4051-4055, 1988
 A:Title: Cloning and sequencing of the cDNA encoding a core protein of the paired hel
 A:Reference number: A30217; MUID:88234557; PMID:3131773
 A:Accession: A30217
 A:Molecule type: mRNA
 A:Residues: 1-44, 103-274, 306-441 <GO2>
 A:Cross-references: GB:J03778; NID:g338684; PIDN:AAA60615.1; PID:g338685
 R:Lee, G.; Neve, R.L.; Kosik, K.S.
 Neuron 2, 1615-1624, 1989
 A:Title: The microtubule binding domain of tau protein.
 A:Reference number: JN0009; MUID:90180482; PMID:2516729

A:Accession: JN0009
A:Molecule type: mRNA
A:Residues: 1-44,103-274,306-441 <LEE>
R:Goedert, M.; Spillantini, M.G.; Potier, M.C.; Ulrich, J.; Crowther, R.A.
EMBO J. 8, 393-399, 1989
A:Title: Cloning and sequencing of the cDNA encoding an isoform of microtubule-associated
A:Reference number: S03796; MUID:89251564; PMID:2498079
A:Accession: S03796
A:Molecule type: mRNA
A:Residues: 1-44,103-441 <G03>
A:Cross-references: EMBL:X14474; NID:g36724; PIDN:CAA32636.1; PID:g36725
R:Andreass, A.; Brown, W.M.; Kosik, K.S.
Biochemistry 31, 10626-10633, 1992
A:Title: Structure and novel exons of the human tau gene.
A:Reference number: S26662; MUID:93041757; PMID:1420178
A:Accession: S26665
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 144-185 <AND>
A:Cross-references: EMBL:X61372; NID:g36718; PID:g36719
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1991
A:Accession: S26666
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 187-274 <AN2>
A:Cross-references: EMBL:X61374; NID:g36722; PID:g36723
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1991
A:Accession: S26662
A:Molecule type: DNA
A:Residues: 371-441 <ANW>
A:Cross-references: EMBL:X61373
R:Jakes, R.; Novak, M.; Davison, M.; Wischik, C.M.
EMBO J. 10, 2725-2729, 1991
A:Title: Identification of 3- and 4-repeat tau isoforms within the PHF in Alzheimer's di
A:Reference number: S17302; MUID:92007714; PMID:1915258
A:Accession: S17302
A:Status: preliminary
A:Molecule type: protein
A:Residues: 268-274,306-395 <JAK>
R:Hasegawa, M.; Morishima-Kawashima, M.; Takio, K.; Suzuki, M.; Titani, K.; Ihara, Y.
J. Biol. Chem. 267, 17047-17054, 1992
A:Title: Protein sequence and mass spectrometric analyses of tau in the Alzheimer's dise
A:Reference number: A43444; MUID:92381012; PMID:1512244
A:Accession: A43444
A:Molecule type: protein
A:Residues: 2-73,103-130,151-180,191-254,260-269;275-290;299-317,322-340;344-347,354-383
A:Experimental source: Alzheimer's disease brain
A:Note: sequence extracted from NCBI backbone (NCBIP:112039)
C:Comment: This heterogeneous protein, which is found predominantly in cells of the nerv
o the core protein of the paired helical filament of Alzheimer's disease.
C:Genetics:
A:Gene: GDB:MAPT
A:Cross-references: GDB:119434; OMIM:157140
A:Map position: 17q21-17q21
C:Superfamily: microtubule-associated protein tau; MAP2/tau repeat homology
C:Keywords: alternative splicing; Alzheimer's disease; duplication; microtubule binding;
F:1-441/Product: microtubule-associated protein tau, long splice form #status predicted
F:1-274,306-441/Product: microtubule-associated protein tau (clone htau39) #status predi
F:1-73,103-441/Product: microtubule-associated protein tau (clone htau34) #status predi
F:1-73,103-274,306-441/Product: microtubule-associated protein tau (clone htau34) #status predi
F:1-44,103-274,306-441/Product: microtubule-associated protein tau (clone htau37) #statu
F:1-44,103-441/Product: microtubule-associated protein tau, fetal #status predi
F:1-44,103-441/Product: microtubule-associated protein tau type II #status predicted <MA
F:252-282/Domain: MAP2/tau repeat homology <WT1>
F:283-313/Domain: MAP2/tau repeat homology <WT2>
F:314-344/Domain: MAP2/tau repeat homology <WT3>
F:345-376/Domain: MAP2/tau repeat homology <WT4>
Query Match 53.4%; Score 2116.5; DB 1; Length 441;
Best Local Similarity 58.2%; Pred. No. 7,9e-90;
Matches 441; Conservative 0; Mismatches 0; Indels 317; Gaps 2;
QY 1 MAEPQEFVEMDHAGTYGLGRKQGGYTMHQDEGDTAGLKESPLQTPEDSGSEPG 60
|||||

Db 1 MAEPQEFVEMDHAGTYGLGRKQGGYTMHQDEGDTAGLKESPLQTPEDSGSEPG 60
QY 61 SETSDAKSTPTAEDVTAPLVDEGAPGKAAAQPHTEIPEGTTAEAGTGPSPLEDEAAG 120
|||||
Db 61 SETSDAKSTPTAEDVTAPLVDEGAPGKAAAQPHTEIPEGTTAEAGTGPSPLEDEAAG 120
|||||
QY 121 HVTQBPESGKVVQEGFLREPGLSHQLMGMPGAPLLPEGPREATRQPSGTGPDTEG 180
|||||
Db 121 HVTQ----- 124
QY 181 GRHAPELLKHOLLGLDHOEGPLKAGGAGKERPGSKKEEVEDRDVDESSPDSPSKASPA 240
Db 125 ----- 124
QY 241 QDGRPPQTAARENTSIPGFAEGAIPLVDFLSKVSTEIPASEPDGSPVGRAGQODAPLE 300
Db 125 ----- 124
QY 301 FTFHVEITPNVQKEQHSEHGLRAAPGAPGEGPEARGPSILGDTKEADLPEPSEKQPA 360
Db 125 ----- 124
QY 361 AAPRGKPVSRVPOLKARVMYSKSDGTGSDDKAKKTSTRSSAKTLKNRPCLSPKLPFGSS 420
|||||
Db 125 -----ARMVSKSDGTGSDDKKA----- 142
QY 421 DPLQPSPPAVCPPEPPSPKIVSVTSRTGSGAKEMKLGADGKTKIATPRGAAPGQK 480
|||||
Db 143 -----KGADGKTKIATPRGAAPGQK 163
QY 481 GOANATRIPTAKTPAPKTPPPSSGPPKSGDRSGYSSPGSGTSGSRRTPSLPTPTPREP 540
|||||
Db 164 GOANATRIPTAKTPAPKTPPPSSGPPKSGDRSGYSSPGSGTSGSRRTPSLPTPTPREP 223
QY 541 KKVAVVTRTPPKSPSSAKSRLQTPVPMPLKLVKSKIGSTENLKHQGGKGVQIINKKLD 600
|||||
Db 224 KKVAVVTRTPPKSPSSAKSRLQTPVPMPLKLVKSKIGSTENLKHQGGKGVQIINKKLD 283
QY 601 LSNVQSKCGSKDNKTHVPGGSGVQIVYKPVDLKSVTSKCGSLGNHIIHKPGGQGVSEK 660
|||||
Db 284 LSNVQSKCGSKDNKTHVPGGSGVQIVYKPVDLKSVTSKCGSLGNHIIHKPGGQGVSEK 343
QY 661 LDFKDRVQSKTGLSDNITHVPGGSGNKKIETHKLTFRENAKAKTDHGAIVYKSPVSGDT 720
|||||
Db 344 LDFKDRVQSKTGLSDNITHVPGGSGNKKIETHKLTFRENAKAKTDHGAIVYKSPVSGDT 403
QY 721 SPRHLNSVSTGSDMDVDSPLATLADSVASLAKQGL 758
|||||
Db 404 SPRHLNSVSTGSDMDVDSPLATLADSVASLAKQGL 441
RESULT 4
JS0306
microtubule-associated protein tau - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 31-Dec-1993
C:Accession: JS0306; A33574
R:Kosik, K.S.; Orcicchio, L.D.; Bakalis, S.; Neve, R.L.
Neuron 2, 1389-1397, 1989
A:Title: Developmentally regulated expression of specific tau sequences.
A:Reference number: JS0306; MUID:90180457; PMID:2560640
A:Accession: JS0306
A:Molecule type: mRNA
A:Residues: 1-432 <KOS>
A:Note: the sequence shown is from adult rat brain
A:Note: the partial sequence from fetal rat brain is lacking residues 266-296; the fe
R:Kanai, Y.; Takemura, R.; Oshima, T.; Mori, H.; Ihara, Y.; Yanagisawa, M.; Masaki, T
J. Cell Biol. 109, 1173-1184, 1989
A:Title: Expression of multiple tau isoforms and microtubule bundle formation in fibr
A:Reference number: A33574; MUID:89359509; PMID:2504728
A:Accession: A33574
A:Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-432 <KAN>

A:Note: a variant lacking residues 63-91 was also found

C:Superfamily: microtubule-associated protein tau; MAP2/tau repeat homology

C:Keywords: alternative splicing; Alzheimer's disease; calmodulin binding; microtubule b

F:243-273/Domain: MAP2/tau repeat homology <MT1>

F:274-304/Domain: MAP2/tau repeat homology <MT2>

F:305-335/Domain: MAP2/tau repeat homology <MT3>

F:336-367/Domain: MAP2/tau repeat homology <MT4>

F:282-313/Disulfide bonds: #status experimental

F:347/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match

Best Local Similarity 46.5%; Score 1845; DB 2; Length 432;

Matches 394; Conservative 15; Mismatches 21; Indels 330; Gaps 4;

QY 1 MAEPRQEFVEMDHACTYGLGRKQGGYTMHQDEGTDAGLKESPLQPTEDGSEEPG 60

Db 1 MAEPRQEFVEMDHACTYGLGRKQGGYTMHQDEGTDAGLKESPLQPTEDGSEEPG 49

QY 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAOAPHTPEIGTTAEAGIGDTPSLDEAAG 120

Db 50 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAOAPHTPEIGTTAEAGIGDTPSLDEAAG 109

QY 121 HVTQPEESGKVVQEGFLREPPGGLSHQLMSGMPGAPLLPEGPREATRQPSGTGPEDETEG 180

Db 110 HVTQ----- 113

QY 181 GRHAPELLKHQLLDLHQBGLPKAGGKERGCKSEEVDEDRDVBESSPQSPKASPA 240

Db 114 ----- 113

QY 241 QDGRPPQTAAREATSIQGPBAEATPLPVDFLSKVSTEIPASEPDPSPVGRAGQDAPLE 300

Db 114 ----- 113

QY 301 FTFHVEITPNVQEQAHSEELHGRAFFGAPGEGPEARGPSLGEDTKADLPPEPKQPA 360

Db 114 ----- 113

QY 361 AAPRGKPSRVPLKARWYKSKDGTSGDDKAKTSTRSSAKTLKNRCLSPKLPPTGSS 420

Db 114 ----- 131

QY 421 DPLIQSPAVCPPEPPSPKHYSSVTSTRTSGGKEMKLGADGKT--KIATPRGAAPP 478

Db 132 ----- 152

QY 479 QKQANATRIAPKTPAPKTPPSGEPKSGDRSGYSSPGTSGSRTPSLTPPTPR 538

Db 153 QKGTSNATRIAPKTPSPKTPPSGEPKSGDRSGYSSPGTSGSRTPSLTPPTPR 212

QY 539 EPKKAIVVTPPKSPSSAKSRLOTAPVPMPLKKNVSKIGSTENLKHQGGGKVQIINK 598

Db 213 EPKKAIVVTPPKSPSSAKSRLOTAPVPMPLKKNVSKIGSTENLKHQGGGKVQIINK 272

QY 599 LDLSNVQSKGSKDNTHKVPVGGGQVQVYKPYDLSKVTSCGSLGNIHKPGGGGVQEVKS 658

Db 273 LDLSNVQSKGSKDNTHKVPVGGGQVQVYKPYDLSKVTSCGSLGNIHKPGGGGVQEVKS 332

QY 659 EKLDFKDRVQSKIGSLDNTHTVPGGNGKKIETHKLTFFRENAKAKTDHGAIEVYKSPVVS 718

Db 333 EKLDFKDRVQSKIGSLDNTHTVPGGNGKKIETHKLTFFRENAKAKTDHGAIEVYKSPVVS 392

QY 719 DTSPRHLNVSTGSDMWDSPLQATLADSVASLAKQGL 758

Db 393 DTSPRHLNVSTGSDMWDSPLQATLADSVASLAKQGL 432

RESULT 5

QRBOT1

N:Contains: microtubule-associated protein tau, form 1 - bovine

microtubule-associated protein tau, form 2

C:Species: Bos primigenius taurus (cattle)

C>Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 22-Jun-1999

C:Accession: A31939; A33914; S04005; A48885; A28173; B33734

R:Himmeler, A.; Drechsel, D.; Kirschner, M.W.; Martin Jr., D.W.

Mol. Cell. Biol. 9, 1381-1388, 1989

A:Title: Tau consists of a set of proteins with repeated C-terminal microtubule-bindin

A:Reference number: A31939; MUID:89261765; PMID:2498649

A:Accession: A31939

A:Molecule type: mRNA

A:Residues: 1-448 <HIM>

A:Cross-references: GB:M26157; NID:g514913; PIDN:AAA30770.1; PID:g514914

R:Iqbal, K.; Grundke-Iqbal, I.; Smith, A.J.; George, L.; Tung, Y.C.; Zaidi, T.

Proc. Natl. Acad. Sci. U.S.A. 86, 5646-5650, 1989

A:Title: Identification and localization of a tau-peptide to paired helical filaments

A:Reference number: A33914; MUID:89315854; PMID:2501795

A:Accession: A33914

A:Molecule type: protein

A:Residues: 28,'A','30-38','IG','41','AP','44','LK' <IQB>

A:Experimental source: brain

A:Note: 40-pro was also found

R:Iqbal, K.; Smith, A.J.; Zaidi, T.; Grundke-Iqbal, I.

FEBS Lett. 248, 87-91, 1989

A:Title: Microtubule-associated protein tau. Identification of a novel peptide from b

A:Reference number: S04005; MUID:89252057; PMID:2498127

A:Accession: S04005

A:Molecule type: protein

A:Residues: 28,'A','30-38','IG','41','AP','44','LK' <IQ2>

A:Experimental source: brain

A:Note: 40-Pro was also found

R:Paudel, H.K.; Lew, J.; Ali, Z.; Wang, J.H.

J. Biol. Chem. 268, 23512-23518, 1993

A:Title: Brain proline-directed protein kinase phosphorylates tau on sites that are a

A:Reference number: A48885; MUID:94043150; PMID:8226879

A:Accession: A48885

A:Molecule type: protein

A:Residues: 'X','203-208','X','210-211','X','213-216;238-241','X','243-247','X','404-410','X',4

A:Experimental source: brain

A:Note: sequence modified after extraction from NCBI backbone

R:Aizawa, H.; Kawasaki, H.; Murofushi, H.; Kotani, S.; Suzuki, K.; Sakai, H.

J. Biol. Chem. 263, 7703-7707, 1988

A:Title: Microtubule-binding domain of Tau proteins.

A:Reference number: A28173; MUID:8827970; PMID:3131325

A:Accession: A28173

A:Molecule type: protein

A:Residues: 205-218,'X','220-223' <A12>

A:Experimental source: brain

C:Superfamily: microtubule-associated protein tau; MAP2/tau repeat homology

C:Keywords: alternative splicing; microtubule binding; phosphoprotein; tandem repeat

F:1-448/Product: microtubule-associated protein tau, form 1 #status predicted <BT43>

F:1-174,193-448/Product: microtubule-associated protein tau, form 2 #status predicted

F:205-223/Region: microtubule binding #status experimental

F:259-289/Domain: MAP2/tau repeat homology <MT1>

F:290-320/Domain: MAP2/tau repeat homology <MT2>

F:321-351/Domain: MAP2/tau repeat homology <MT3>

F:352-383/Domain: MAP2/tau repeat homology <MT4>

F:202,209,242,248,411/Binding site: phosphate (Ser) (covalent) (by proline-directed k

F:212/Binding site: phosphate (Thr) (covalent) (by proline-directed kinase) #status e

Query Match 46.5%; Score 1845; DB 1; Length 448;

Best Local Similarity 51.7%; Pred. No. 1.9e-77;

Matches 401; Conservative 9; Mismatches 20; Indels 346; Gaps 6;

QY 1 MAEPRQEFVEMDHACTYGLGRKQGGYTMHQDEGTDAGLKESPLQPTEDGSEEPG 60

Db 1 MAEPRQEFVEMDHACTYGLGRKQGGYTMHQDEGTDAGLKESPLQPTEDGSEEPG 49

QY 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAOAPHTPEIGTTAEAGIGDTPSLDEAAG 120

Db 50 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAOAPHTPEIGTTAEAGIGDTPSLDEAAG 109

QY 121 HVTQPEESGKVVQEGFLREPPGGLSHQLMSGMPGAPLLPEGPREATRQPSGTGPEDETEG 180

Db 110 HVTQ----- 113

QY 181 GRHAPPELLKHQLLDGDLHQEGPLKAGGKERPGSKEEVEDRVDSESSPQSPKASPA 240
 Db 114 -----
 QY 241 QDGRPPQTAAREATSIPGFPABGA1PLPVDFLSKVSVEIPASEPDGPGSVGRAKQDAPLE 300
 Db 114 -----
 QY 301 FTFHVEITPNVQEQAHSEHLGRAAFPCAPGEGPEARGPSLGEDTKADLPPESEKQPA 360
 Db 114 -----
 QY 361 AAPRGKPSRVPLQKARMYSKSDGTGDDKAKTSTRSSAKTLKRNPCLSPLKPTPGSS 420
 Db 114 -----
 QY 421 DPLIQSSPAVCPPEPPSPKHVSSTRTGSSGAKEMKLGADGK--TKIATPRGAAPP 478
 Db 132 -----
 QY 479 QKGOANATRIAPKTPAKTPPSS-----GPPKSGDRSGSYSSPGSPGT 522
 Db 153 QKGOANATRIAPKTPPTPKTSPATMOVKPPAGAKSERGESGSGDRSGSYSSPGSPGT 212
 QY 523 PGSRRTPSLPTPTREPCKVAVRTPPKSPSAKSRLQTAPVPMPLKNVYSKIGSTEN 582
 Db 213 PGSRRTPSLPTPTREPCKVAVRTPPKSPSAKSRLQAAPCPMDLKNVYSKIGSTEN 272
 QY 583 LKHQPGGKVQIINKKLDLSNVQSCGSKDNKIKHVPGGSGVQIVYKPVLDLSKVTSCGSL 642
 Db 273 LKHQPGGKVQIINKKLDLSNVQSCGSKDNKIKHVPGGSGVQIVYKPVLDLSKVTSCGSL 332
 QY 643 GNTHKPGGQVEVSEKLDKDFRVQSKIGSLDNITHTHVPGGNKKIETHKLTFRENAKAK 702
 Db 333 GNTHKPGGQVEVSEKLDKDFRVQSKIGSLDNITHTHVPGGNKKIETHKLTFRENAKAK 392
 QY 703 TDHGAIEIVKSPVSGDTSRPHLSNVSTGSDMDVDSPLQATLADEVSAKAKQGL 758
 Db 393 TDHGAIEIVKSPVSGDTSRPHLSNVSTGSDMDVDSPLQATLADEVSAKAKQGL 448
 RESULT 6
 QRB02
 N:Contains: microtubule-associated protein tau, form 3 - bovine
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 31-Mar-1996
 R:Himmeler, A.; Drechsel, D.; Kirschner, M.W.; Martin Jr., D.W.
 Mol. Cell. Biol. 9, 1381-1388, 1989
 A:Title: Tau consists of a set of proteins with repeated C-terminal microtubule-binding
 A:Reference number: A31939; MUID:89261765; PMID:2498649
 A:Accession: B31939
 A:Molecule type: mRNA
 A:Residues: 1-402 <HIM>
 A:Cross-references: GB:M26157; GB:M26158
 R:Paudel, H.K.; Lew, J.; Ali, Z.; Wang, J.H.
 J. Biol. Chem. 268, 23512-23518, 1993
 A:Title: Brain proline-directed protein kinase phosphorylates tau on sites that are abno
 A:Reference number: A48885; MUID:94043150; PMID:8226879
 A:Accession: A48885
 A:Molecule type: protein
 A:Residues: 'X', 157-162, 'X', 164-165, 'X', 167-170, 192-195, 'X', 197-201, 'X', 358-364, 'X', 366-
 A:Experimental source: brain
 A>Note: sequence modified after extraction from NCBI backbone
 R:Aizawa, H.; Kawasaki, H.; Murofushi, H.; Kotani, S.; Suzuki, K.; Sakai, H.
 J. Biol. Chem. 263, 7703-7707, 1988
 A:Title: Microtubule-binding domain of Tau proteins.
 A:Reference number: A28173; MUID:88227970; PMID:3131325
 A:Accession: A28173
 A:Molecule type: protein
 A:Residues: 159-172, 'X', 174-177 <AIZ>

A:Experimental source: brain
 C:Superfamily: microtubule-associated protein tau; MAP2/tau repeat homology
 C:Keywords: alternative splicing; microtubule binding; phosphoprotein; tandem repeat
 F:1-402/Product: microtubule-associated protein tau, form 3 #status predicted <Bf4>
 F:1-234,297-402/Product: microtubule-associated protein tau, form 5 #status predicted
 F:101-402/Product: microtubule-associated protein tau, form 4 #status predicted
 F:159-177/Region: microtubule binding #status experimental
 F:213-243/Domain: MAP2/tau repeat homology <MT1>
 F:244-274/Domain: MAP2/tau repeat homology <MT2>
 F:275-305/Domain: MAP2/tau repeat homology <MT3>
 F:306-337/Domain: MAP2/tau repeat homology <MT4>
 F:156,163,196,202,365/Binding site: phosphate (Ser) (covalent) (by proline-directed k
 F:166/Binding site: phosphate (Thr) (covalent) (by proline-directed kinase) #status p

Query Match 44.6%; Score 1770; DB 1; Length 402;
 Best Local Similarity 94.8%; Pred. No. 4.5e-74;
 Matches 347; Conservative 4; Mismatches 12; Indels 4; Gaps 2;

QY 394 KTSTRSSAKTLKRNPCLSPLPTPGSSDPLIQSPSPKSPKSVTSRTGSSG 453
 Db 38 KPSTPSSAKTLKRNPCLSPKRPTGSSDPLIKPSPSPKSPKSVTSRTGSSG 97
 QY 454 AKEMKLGADGK--TKIATPRGAAPPQKGOANATRIAPKTPPAPKTPPSSGEPKSGDR 511
 Db 98 AKEMKVGADGKPGTKIATPRGAAPPQKGOANATRIAPKTPPPTKTP--GESGKSGDR 155
 QY 512 SGYSSPSGPTPGSRRTPSLPTPTREPCKVAVRTPPKSPSAKSRLQTAPVPMPLDK 571
 Db 156 SGYSSPSGPTPGSRRTPSLPTPTREPCKVAVRTPPKSPSAKSRLQAAPCPMDLKNVYSKIGSTEN 215
 QY 572 NVYSKIGSTENLKHQPGGKVQIINKKLDLSNVQSCGSKDNKIKHVPGGSGVQIVYKPVLD 631
 Db 216 NVYSKIGSTENLKHQPGGKVQIINKKLDLSNVQSCGSKDNKIKHVPGGSGVQIVYKPVLD 275
 QY 632 LSKVTSKCGSLGNTHHKPGGQVEVSEKLDKDFRVQSKIGSLDNITHTHVPGGNKKIETH 691
 Db 276 LSKVTSKCGSLGNTHHKPGGQVEVSEKLDKDFRVQSKIGSLDNITHTHVPGGNKKIETH 335
 QY 692 KLTFRENAKAKTDHGAIEIVKSPVSGDTSRPHLSNVSTGSDMDVDSPLQATLADEVSA 751
 Db 336 KLTFRENAKAKTDHGAIEIVKSPVSGDTSRPHLSNVSTGSDMDVDSPLQATLADEVSA 395
 QY 752 SLAKQGL 758
 Db 396 SLAKQGL 402

RESULT 7
 S46264
 microtubule-associated protein - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 13-Aug-1999
 C:Accession: S46264
 R:Sadot, E.; Marx, R.; Barg, J.; Behar, L.; Ginzburg, I.
 J. Mol. Biol. 241, 325-331, 1994
 A:Title: Complete sequence of 3'-untranslated region of tau from rat central nervous
 A:Reference number: S46264; MUID:94334997; PMID:8057376
 A:Accession: S46264
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-374 <SAD>
 A:Cross-references: EMBL:X79321; NID:g517393; PIDN:CAA55889.1; PID:g517394
 C:Superfamily: microtubule-associated protein tau; MAP2/tau repeat homology
 F:185-215/Domain: MAP2/tau repeat homology <MT1>
 F:216-246/Domain: MAP2/tau repeat homology <MT2>
 F:247-277/Domain: MAP2/tau repeat homology <MT3>
 F:278-309/Domain: MAP2/tau repeat homology <MT4>

Query Match 39.1%; Score 1552; DB 2; Length 374;
 Best Local Similarity 45.3%; Pred. No. 3.6e-64;
 Matches 344; Conservative 14; Mismatches 14; Indels 388; Gaps 5;

QY 1 MAEPQFEVMEHDHAGTYGLGDRKDQGGYTMHQDEGDTAGLKESPLQPTPEDGSEEP 60

```
|||||: ||| | | ||| |||| | |||
Db 1 MAEPQEFDTMEDQA-----GDYTLQDQEGMDHGLK----- 33
QY 61 SETSDAKSTPTAEDVTAPLVDGAPGKQAAQPHTEIEGTTAREAGIGDTPSLEDAAG 120
Db 34 -----AEEAGIGDTPNMEDQAAG 51
QY 121 HVTQEPESKVVQEGFLREPGLSHQMSGMPGAPLLPEGPREATRQPSGTGPEDTEG 180
Db 52 HVTQ----- 55
QY 181 GRHAPELLKHQLLDGDLHQEGPLKAGGKPERGSKKEEVEDRDESSPQSPSKASPA 240
Db 56 ----- 55
QY 241 ODGRPPQTAAREATSI PGFPAEGAIPLPVDFLSKVSTEIPASEPDGSPVGRAGGADAPLE 300
Db 56 ----- 55
QY 301 FTFHVEITPNVQEQAHSEHILGRAAFPAGPGEPEARGPSILGEDTKHEADLPSEKOPA 360
Db 56 ----- 55
QY 361 AAPRGKPSVRPQLKARMVSKSGDGTGDDKAKTSTRSSAKTLKNRCLSPKLPDGGSS 420
Db 56 -----ARVAGVSKDRTGNDEKKA----- 73
QY 421 DPLIOPSSPAVCEPPSPKPVSVTSRTGSSGAKEMKLKGADGKT--KIATPRGAAPP 478
Db 74 -----KGADGKTGAKIATPRGAATPG 94
QY 479 QKQANATRIAPKTPPSSGEPKSGDRSGYSPGTPGSRSPGTPSLPTPTPR 538
Db 95 QKGTNATRIAPKTPPSPKTPSGEPKSGERSGYSPGTPGSRSPGTPSLPTPTPR 154
QY 539 EPKKAIVVTPPKSPSSAKSRLOTPAPVMPDLKNVSKIGSTENLKHQPGGKQVQIINK 598
Db 155 EPKKAIVVTPPKSPSSAKSRLOTPAPVMPDLKNVRSKIGSTENLKHQPGGKQVQIINK 214
QY 599 LQLSNVQSKGSKDNIKHVPGGSGVQIVKPVLDLSKVTSCGSLGNIHHKPGGQVEVKS 658
Db 215 LQLSNVQSKGSKDNIKHVPGGSGVQIVKPVLDLSKVTSCGSLGNIHHKPGGQVEVKS 274
QY 659 EKLDFKDRVQSKIGSLDNITHVPGGKNIETHKLTFRENAKAKTDHGAEIVKSPVYSG 718
Db 275 EKLDFKDRVQSKIGSLDNITHVPGGKNIETHKLTFRENAKAKTDHGAEIVKSPVYSG 334
QY 719 DTSPRLSNVSTGSDIMVDSPOLATLADEVSAKQGL 758
Db 335 DTSPRLSNVSTGSDIMVDSPOLATLADEVSAKQGL 374

RESULT 8
QRHUT2
microtubule-associated protein tau, fetal (clone p18) - human
C:Species: Homo sapiens (man)
C>Date: 30-Jun-1990 #sequence_revision 30-Jun-1990 #text_change 02-Sep-1997
C:Accession: P00001
R:Lee, G.; Neve, R.L.; Kosik, K.S.
Neuron 2, 1615-1624, 1989
A:Title: The microtubule binding domain of tau protein.
A:Reference number: JN0009; MUID:90180482; PMID:2516729
A:Accession: P00001
A:Molecule type: mRNA
A:Residues: 1-316 <LEE>
A>Note: this sequence differs from a previously reported fetal tau protein sequence only
C:Genetics:
A:Gene: GDB:MAPT; MTBT1
A:Cross-References: GDB:119434; OMIM:157140
A:Map position: 17q21-17q21
C:Superfamily: microtubule-associated protein tau; MAP2/tau repeat homology
C:Keywords: alternative splicing; Alzheimer's disease; duplication; microtubule binding;
F:158-188/Domain: MAP2/tau repeat homology <MT1>
```

```
F:189-219/Domain: MAP2/tau repeat homology <MT2>
F:220-251/Domain: MAP2/tau repeat homology <MT3>

Query Match 35.7%; Score 1415.5; DB 1; Length 316;
Best Local Similarity 74.5%; Pred. No. 5.1e-58;
Matches 286; Conservative 1; Mismatches 0; Indels 97; Gaps 2;

QY 375 KARMVSKSGDGTGDDKKAKTSTRSSAKTLKNRCLSPKLPDGGSSDPLQPSPAVCP 434
Db 30 QARVSKSGDGTGDDKKA----- 48
QY 435 PPSPKHVSVSVTSRTSGSGAKEMKLKGADGKTKIATPRGAAPPQKQANATRIAPKTPP 494
Db 49 -----KGADGKTKIATPRGAAPPQKQANATRIAPKTPP 83
QY 495 APKTPSSGEPKSGDRSGYSPGTPGSRSPGTPSLPTPTPREPKKAVVTPPKSPS 554
Db 84 APKTPSSGEPKSGDRSGYSPGTPGSRSPGTPSLPTPTPREPKKAVVTPPKSPS 143
QY 555 SAKSRLQATAPVMPDLKNVSKIGSTENLKHQPGGKQVQIINKLQLSNVQSKGSKDNI 614
Db 144 SAKSRLQATAPVMPDLKNVSKIGSTENLKHQPGGK----- 180
QY 615 KHPVGGSGVQIVKPVLDLSKVTSCGSLGNIHHKPGGQVEVSEKLDKDFKRVQSKIGSL 674
Db 181 -----VOIVKPVLDLSKVTSCGSLGNIHHKPGGQVEVSEKLDKDFKRVQSKIGSL 232
QY 675 DNITHVPGGKNIETHKLTFRENAKAKTDHGAEIVKSPVVS GDTSPRLSNVSTGSI 734
Db 233 DNITHVPGGKNIETHKLTFRENAKAKTDHGAEIVKSPVVS GDTSPRLSNVSTGSI 292
QY 735 DMVDSPOLATLADEVSAKQGL 758
Db 293 DMVDSPOLATLADEVSAKQGL 316

RESULT 9
B28820
microtubule-associated protein tau type 2 - mouse
C:Species: Mus musculus (house mouse)
C>Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 13-Aug-1999
C:Accession: B28820
R:Lee, G.; Cowan, N.; Kirschner, M.
Science 239, 285-288, 1988
A:Title: The primary structure and heterogeneity of tau protein from mouse brain.
A:Reference number: A94298; MUID:88099510; PMID:3122323
A:Accession: B28820
A:Molecule type: mRNA
A:Residues: 1-341 <LEE>
A:Cross-References: GB:M18775; NID:q201114; PIDN:AAA0165.1; PID:q201115
C:Superfamily: microtubule-associated protein tau; MAP2/tau repeat homology
C:Keywords: alternative splicing; microtubule binding; tandem repeat
F:183-213/Domain: MAP2/tau repeat homology <MT1>
F:214-244/Domain: MAP2/tau repeat homology <MT2>
F:245-276/Domain: MAP2/tau repeat homology <MT3>

Query Match 34.2%; Score 1354.5; DB 2; Length 341;
Best Local Similarity 41.2%; Pred. No. 3.3e-55;
Matches 313; Conservative 13; Mismatches 13; Indels 421; Gaps 7;

QY 1 MAEPQEFYEMDHAGTYGIGDRKDGQGYTMHQDQEGDTDAGLKESPLQPTPTEDGSEEPG 60
Db 1 MADPRQEFDTMEDHA-----GDYTLQDQEGMDHGLK----- 33
QY 61 SETSDAKSTPTAEDVTAPLVDGAPGKQAAQPHTEIEGTTAREAGIGDTPSLEDAAG 120
Db 34 -----AEEAGIGDTPNQEDQAAG 51
QY 121 HVTQEPESKVVQEGFLREPGLSHQMSGMPGAPLLPEGPREATRQPSGTGPEDTEG 180
Db 52 HVTQ----- 55
QY 181 GRHAPELLKHQLLDGDLHQEGPLKAGGKPERGSKKEEVEDRDESSPQSPSKASPA 240
```

Db 56 ----- 55
QY 241 QDGRPPQTAAREATSI PGPAEGAIPLPVDFLSKVSTEIPASEPDGPSVGRAGQDAPLE 300
Db 56 ----- 59
QY 301 FTFHVEITPNVQKEAHSEHLGRAAFPAGPGEPEARGLGEDTKADLPPEPSEKQPA 360
Db 60 ----- 59
QY 361 AAPRGKPVSRVQPKARVMVSKSDGTGSDDKKAKTSTRSSAKTLKNRCLSPKLPPTGSS 420
Db 60 ----- 71
QY 421 DPLIQSSPAVCPPEPPSPKHVSSVTSRTGSSGAKEMKLKGADGKT--KIATPRGAAPP 478
Db 72 ----- 92
QY 479 QKQANATRIPAKTPAPKTPPSSGEPKSGDRSGYSSPGTSGSRSTPSLTPPTR 538
Db 93 QKTSNATRIPAKTPSPKTPPGSGEPKSGERSGYSSPGTSGSRSTPSLTPPTR 152
QY 539 EPKKVAVVRTPPKSPSASRSRLQTAAPVMPDLKNVSKIGSTENLKHQGGGKVQIINKK 598
Db 153 EPKKVAVVRTPPKSPSASRSRLQTAAPVMPDLKNVSKIGSTENLKHQGGGK----- 205
QY 599 LDLSNVQSKGSDKNIKHVPGGGSGVQIYKVPDLSKVTSCGSLGNIHHKPGGGOVEVKS 658
Db 206 ----- 241
QY 659 EKLDFKDRVQSKIGSLDNITHVPGGNGKKIETHKLTFRENAKAKTDHGAIEIVKSPVVS 718
Db 242 EKLDFKDRVQSKIGSLDNITHVPGGNGKKIETHKLTFRENAKAKTDHGAIEIVKSPVVS 301
QY 719 DTSPRHLSNVSTGSDIMVDSPOLATLADEVASLAKQGL 758
Db 302 DTSPRHLSNVSTGSDIMVDSPOLATLADEVASLAKQGL 341
RESULT 10
A28820
microtubule-associated protein tau type 1 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 13-Aug-1999
C:Accession: A28820
R:Lee, G.; Cowan, N.; Kirschner, M.
Science 239, 285-288, 1988
A:Title: The primary structure and heterogeneity of tau protein from mouse brain.
A:Reference number: A94298; MUID:88099510; PMID:3122323
A:Accession: A28820
A:Molecule type: mRNA
A:Residues: 1-364 <LE>
A:Cross-references: GB:M18776; NID:g201116; PIDN:AA40166.1; PID:g201117
C:Superfamily: microtubule-associated protein tau; MAP2/tau repeat homology
C:Keywords: alternative splicing; microtubule binding; tandem repeat
F:183-213/Domain: MAP2/tau repeat homology <MT1>
F:214-244/Domain: MAP2/tau repeat homology <MT2>
F:245-276/Domain: MAP2/tau repeat homology <MT3>
Query Watch 34.1%; Score 1350.5; DB 2; Length 364;
Best Local Similarity 41.1%; Pred. No. 5.4e-55;
Matches 312; Conservative 13; Mismatches 13; Indels 421; Gaps 7;
QY 1 MAP2PEFEFVMEHAGTYGLGDRKDGQGYTMHQDQEGDNDAGLKSPLOTPTEDGSEEPG 60
Db 1 MADPREFEFTMDHA-----GDYTLQLQDQEGDMDHGLK----- 33
QY 61 SETSDAKSTPTABDVTAPLVDEGAPCKQAQAQPHTEIPGTTAEAGIGDTPSLEDEAAG 120
Db 34 -----AEEAGIGDTPNQDAQAG 51
QY 121 HVTQPESGKVQVQEGFLREFGPPGLSHQLMSGNPGAPLLPEGPREATROPSTGTPEDTGG 180

Db 52 HVTQ----- 55
QY 181 GRHAPELLKHQLLDGLHQEGBPPLKGAGGKERPGSKEEVEDRDVDDESSPDSPSKASPA 240
Db 56 ----- 55
QY 241 QDGRPPQTAAREATSI PGPAEGAIPLPVDFLSKVSTEIPASEPDGPSVGRAGQDAPLE 300
Db 56 ----- 59
QY 301 FTFHVEITPNVQKEAHSEHLGRAAFPAGPGEPEARGLGEDTKADLPPEPSEKQPA 360
Db 60 ----- 59
QY 361 AAPRGKPVSRVQPKARVMVSKSDGTGSDDKKAKTSTRSSAKTLKNRCLSPKLPPTGSS 420
Db 60 ----- 71
QY 421 DPLIQSSPAVCPPEPPSPKHVSSVTSRTGSSGAKEMKLKGADGKT--KIATPRGAAPP 478
Db 72 ----- 92
QY 479 QKQANATRIPAKTPAPKTPPSSGEPKSGDRSGYSSPGTSGSRSTPSLTPPTR 538
Db 93 QKTSNATRIPAKTPSPKTPPGSGEPKSGERSGYSSPGTSGSRSTPSLTPPTR 152
QY 539 EPKKVAVVRTPPKSPSASRSRLQTAAPVMPDLKNVSKIGSTENLKHQGGGKVQIINKK 598
Db 153 EPKKVAVVRTPPKSPSASRSRLQTAAPVMPDLKNVSKIGSTENLKHQGGGK----- 205
QY 599 LDLSNVQSKGSDKNIKHVPGGGSGVQIYKVPDLSKVTSCGSLGNIHHKPGGGOVEVKS 658
Db 206 ----- 241
QY 659 EKLDFKDRVQSKIGSLDNITHVPGGNGKKIETHKLTFRENAKAKTDHGAIEIVKSPVVS 718
Db 242 EKLDFKDRVQSKIGSLDNITHVPGGNGKKIETHKLTFRENAKAKTDHGAIEIVKSPVVS 301
QY 719 DTSPRHLSNVSTGSDIMVDSPOLATLADEVASLAKQGL 757
Db 302 DTSPRHLSNVSTGSDIMVDSPOLATLADEVASLAKQGL 340
RESULT 11
GRHUMT
microtubule-associated protein 2, splice form MAP-2b - human
N:Alternate names: MAP2
C:Species: Homo sapiens (man)
C:Date: 31-Mar-1989 #sequence_revision 12-Apr-1996 #text_change 21-Jul-2000
C:Accession: I53693; A61085; PL0024; S34131
R:Albala, J.S.; Kalcheva, N.; Shafit-Zagardo, B.
Gene 136, 377-378, 1993
A:Title: Characterization of the transcripts encoding two isoforms of human microtubule-associated protein 2.
A:Reference number: I53693; MUID:94124038; PMID:8294038
A:Accession: I53693
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1824 <ALB>
A:Cross-references: EMBL:Z21958; GB:L12563; NID:gl850616; PIDN:AAB48098.1; PID:gl8506
R:Dammerman, M.; Yen, S.H.; Shafit-Zagardo, B.
J. Neurosci. Res. 24, 487-495, 1989
A:Title: Sequence of a human MAP-2 region sharing epitopes with Alzheimer neurofibril
A:Reference number: A61085; MUID:90096190; PMID:2481044
A:Accession: A61085
A>Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 77-645 <DAM>
R:Kosik, K.S.; Orecchio, L.D.; Bakalis, S.; Duffy, L.; Neve, R.L.
J. Neurochem. 51, 587-598, 1988
A:Title: Partial sequence of MAP2 in the region of a shared epitope with Alzheimer ne
A:Reference number: PL0024; MUID:88274407; PMID:2455776
A:Accession: PL0024

A:Molecule type: mRNA
A:Residues: 489-1538 <KOS>
A:Cross-references: GB:M25668; NID:g187380; PIDN:AAA59552.1; PID:g187381
C:Comment: Microtubule-associated proteins are a complex group consisting of the high molecular weight proteins that they may stabilize the microtubules against depolymerization.
C:Genetics:
A:Gene: GDB:MAP2
A:Cross-references: GDB:l18836; OMIM:157130
A:Map position: 2q34-2q35
A:Superfamily: microtubule-associated protein MAP2b; MAP2/tau repeat homology
C:Keywords: alternative splicing; microtubule binding; phosphoprotein; tandem repeat
F:1452-1463/Region: microtubule binding #status predicted
F:1666-1696/Domain: MAP2/tau repeat homology <MT1>
F:1697-1727/Domain: MAP2/tau repeat homology <MT2>
F:1728-1759/Domain: MAP2/tau repeat homology <MT3>
F:657,958,1064,1250,1436,1503/Binding site: phosphate (Thr) (covalent) #status predicted
F:697,817,829,1320,1417,1542,1551/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 20.4%; Score 812.5; DB 1; Length 1824;
Best Local Similarity 29.2%; Pred. No. 1e-29;
Matches 284; Conservative 120; Mismatches 279; Indels 291; Gaps 39;

QY 5 ROFEFVEMDHAGTYGLGDRKDGKGYTMHODQSGDTDAGLKESPLQPTEDGSEEPGSETS 64
Db 922 KDEFVS-DKEASAHSGDK---SGUSKEFDQKKAN-----DRLDTVLEK-SEEHADSK 971
QY 65 DAKSTPTAEDVTAPLVDEGAPGKQAAPHTPEIGTTAEAA--GIGDTPLSEDEAAGHV 122
Db 972 HAKKTEAGD-ELETFLGLVTEYEQALAK-DLSIPTDASSEKAEGKLSVPEI----- 1021
QY 123 TORPESGKVVQGF-----LREPGPGLSHQLMSGM-----PGAPLLPGPREA 166
Db 1022 -AEVPSKKVEQGLDFAVGQGLDVKISDFG-----QMASGLNIDRRATELKLKLEATQDM 1074
QY 167 TROPSTGTPEDT-----EGGRHAPPELLKHOLLGD--LHOEGPPLK--GAGGK 209
Db 1075 T--PSSKAPQEAADAFMGVSGHMKETKVSETEVKOKVAKPDLVHQAEVADKEESYESSGE 1132
QY 210 ERPGSKEEVEDRDVDESSPODS-----PP--SKASPAQDCRP----- 245
Db 1133 HESLTWESLKADGKKETSPESLIQDETAVKLVSVEIPCPPAVSEADLATDERADVQMEF 1192
QY 246 ---PQTAAREATSIQFPAEGAIPVDFLSKVSTEIPASEPDSVGRAGQDAPLFT 302
Db 1193 IOGPKESKETDITSTPSDVAEPLHETIVSE-PAEIOSEEE---TEAAGEYDKLLFR 1247
QY 303 F-----HVEITPNVQK-----EQAHSEHHLGRAFP 329
Db 1248 SDTLQITDLGVSGAREEFVETCPSEHKGVIESVVTIEDDFTTVQTTTDEGESGSHVRF 1307
QY 330 APGEFPEARGPSLGEDTKREADLPEPSEKO-----PAAAPRGKPVSRVPQLKARMS 380
Db 1308 AALEQPEVERRPSPHDEEFEEVEEAQAEPKDGSPAPASPEREEVA-LSEYKTEIYD 1366
QY 381 KSKDGTGSD-----KKATSTRSS-AKTLKNRP 408
Db 1367 DYKDETTIDDSIMDADSLWVDQDDRSIMTQLETIPKEEKAERARRSLKHKRKEKP 1426
QY 409 C-----LSPKLPPTGSSDPL-----IOPSSPA--VCPPEP 436
Db 1427 FKTGRGRISTPERKVAKEPSTVSDEVRKKAIVYKKAELAKKTEVQAHSPSRKFIKPA 1486
QY 437 ---SSPKHVSYTSRTGSSGA-----KEMKLGADGKTKIATPRGAAP-----PG 478
Db 1487 IKYTRPRLSCVKRKTAAAGGESALAPSVFKQAKDKVSDGVTKSPKRSLLPRPSSILPP 1546
QY 479 QKG-----QANATRIKAPT-----PAPKTPPSS---GEPKSG 509
Db 1547 RRGVSGDRDENSFSLNSSISSARTRSEPIRRAGKSGTSTPTTPPGSTAITPGTPPSS 1606
QY 510 DRSGSYSGSGPTGSGRSRSTPLSTPP-----TREPKKVAVVTRTPKSPSSAKSLQATP 564
Db 1607 SR-----TPGTGTPP-SYPTPTPTGTPKSAIILVPSEKKVIAIIRTPKSPGLTPKQLRLIN 1661

QY 565 VPMPLKNVSKIGSTENLKHQPGGKGVQIINKLKLDSNVQSKGSKDNIKHVPGGGSVQ 624
Db 1662 QPLPDLKNVSKIGSTDNIKYQPKGGQVIVTKKIDLSH----- 1700
QY 625 IVYXPVDSLKVTSKCGSLGNHHPKPGGQVEVSEKLDKDRVQSKIGSLDNTHVPPGGG 684
Db 1701 -----VTSKCGSLKNIRHRPGRGVKIESVKLDKFEKAQAKVSGSLDNAHHVPPGG 1750
QY 685 NKTIETHKLTFRENAKAKTDHGAIEIVYKSPVSGDTPSRHLSNVSSSTGSDMDVDSPOLAT 744
Db 1751 NVKIDSQKLNPREHAKARVDHGAIIITQSGRSSVASPRRLSNVSSSGSINLLESOLAT 1810
QY 745 LADEVASLASLAKQGL 758
Db 1811 LAEDVTAALAKQGL 1824

RESULT 12
A40115
microtubule-associated protein MAP2 - mouse
N:Alternate names: MAP2
C:Species: Mus musculus (house mouse)
C:Date: 27-Mar-1992 #sequence_revision 27-Mar-1992 #text_change 13-Aug-1999
C:Accession: A40115; S06467
R:Lewis, S.A.; Wang, D.; Cowan, N.J.
Science 242, 936-939, 1988
A:Title: Microtubule-associated protein MAP2 shares a microtubule binding motif with A:Reference number: A40115; MUID:89043973; PMID:3142041
A:Accession: A40115
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1828 <LEW>
A:Cross-references: GB:M21041; NID:g199022; PIDN:AAA39490.1; PID:g199023
R:Wang, D.; Lewis, S.A.; Cowan, N.J.
Nucleic Acids Res. 16, 11369-11370, 1988
A:Title: Complete sequence of a cDNA encoding mouse MAP2.
A:Reference number: S06467; MUID:89083571; PMID:3205744
A:Accession: S06467
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-1828 <NAN>
A:Cross-references: GB:M21041; NID:g199022; PIDN:AAA39490.1; PID:g199023
C:Superfamily: microtubule-associated protein MAP2b; MAP2/tau repeat homology
C:Keywords: microtubule binding; tandem repeat
F:1670-1700/Domain: MAP2/tau repeat homology <MT1>
F:1701-1731/Domain: MAP2/tau repeat homology <MT2>
F:1732-1763/Domain: MAP2/tau repeat homology <MT3>

Query Match 20.4%; Score 807; DB 2; Length 1828;
Best Local Similarity 28.3%; Pred. No. 1.8e-29;
Matches 266; Conservative 127; Mismatches 274; Indels 274; Gaps 33;

QY 3 EPROFEFVEMDHAGTYGLGDRKDGK-----YTHHQDQSGDTDAGLKESPLQPTEDGSE- 57
Db 977 EHAKSEMGKGVFLGIGITYDQASTKEKITTTKDTSPKTEKGLSVPEVAEVEPTTKA 1036
QY 58 EPGETSDAKSTPTAEDVTAPLVDEGAPGKQA-----AAQPHTEIPEGTAAEAGIGDTPS 113
Db 1037 DOGLDFPAAKTAEPSSQDIKVSDFGQMASGMNVNAGKAEIKFEVAQBELTSLSSAPQADS 1096
QY 114 LEDBAAGHVTOEPESKVVQEGFLREPGPGLSHQLMSGMPGAPLLPEGPREATRPSQST 173
Db 1097 FMGVESGHI--KEGGKVNTEVKETVKPDLVHQ-----EAVDKESYESSG- 1141
QY 174 GPDETEGRHAPPELLKQLLDGLHQEGPPLKGAGGKRPESKEEVEDRDVDESSPODS- 232
Db 1142 -----EHESL-----TMESLKPDECKKETSPTSLS 1166
QY 233 -----PPSKASPAQDCRP-----PQTAAREATSIQFPAEGAIP 266
Db 1167 IQDEVALKLSVEIPCPPPVSEADLSTDEKEVQMEFQTLKPEESTETPDIPAIPASVOTQP 1226

Db 1317 -PVQPEERRPYDHEELEVLMAAEQAEPKDGSPDAPATPEKEEVEPSEYKTYTIDDDYK 1375
 QY 314 EOAHSSEHLGRA-----APGAPGEGPEARGSLGEDITHEADLPE 353
 Db 1376 DETTIDOSIMDASLWVDOTQDDRSIITEQLETTIPKEERAKEKARRPSLEKHKE----- 1430
 QY 354 PSEKQPAAPAPKPVSRVPQLKARMSKSDGTGSD-----KKAKTSTR 398
 Db 1431 ----KPEKTGRGR--ISTPE---RKVAKKEPSTVSROEVRKKAVYKKAELAKSEVQAH 1481
 QY 399 SSAKTLNRRCLSPKLPFTPQSSDPLIOSSPAVCPPEPPSPKHWSSVTSRT----- 449
 Db 1482 SPSRKLILKPAIKYTRPT-----HLSCVKRKRTTATSGESAQ 1517
 QY 450 GSSGAKEMKLGADGKTKIATPRGAAP-----PGQKG-----QANATRIPAKTPPAKPT 498
 Db 1518 APSAFKAQAKVDYDGTITKSPKRSLSRPSSILPPRGVSGDRENSFSLNSSISSAART 1577
 QY 499 PPSGGEPPKSGDRSGSYSPGSPG-----TPGOS-----RSRTPSLTPPT--REP----- 540
 Db 1578 TRS--EPIRRAGKSGTPTTPGTAITPGTPPPSYSSRTPGTPTSPVPTPGPKFGIL 1635
 QY 541 ----KKVAVVTRTPKSPSSAKSRLQTPAPVMPDLKNVYSKIGSTENLKHQPGGKVVQIIN 596
 Db 1636 VPSEKKVAIIRTPPKSPATPK-QURLINQPLDPLUKNVYSKIGSTDNIKYQPKGQOVQIVT 1694
 QY 597 KKLDLSNVQSKGSKDNIKHVPGGSGVQIVYKVPVDLSKVTSCGSLGNIHHKPGGGQVEV 656
 Db 1695 KKIDLSH-----VTSKCGSLNRIHRPGGGRVKI 1723
 QY 657 KSEKLDPKDRVQSKIGSLDNIHTHVPGGNNKIEHTKITTFRENAKAKTDHGAIEVYKSPVV 716
 Db 1724 ESKVLDPEKRAQKAVGSLDNAHHVPGGNNKIDSQKLNFRHAKARVDHGAIEITQSPSR 1783
 QY 717 SGDTSPRHLSNVSTGSDIMVDSPOLATLADEVASLAKOGL 758
 Db 1784 SSVASPRLLNVSSSGSINLLESPOLATLAEDVTAALAKOGL 1825
 RESULT 15
 167793
 microtubule-associated protein 2, splice form MAP-2c - human
 C:Species: Homo sapiens (man)
 C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 21-Jul-2000
 C:Accession: I67793
 R:Albala, J.S.; Kalcheva, N.; Shafit-Zagardo, B.
 Gene 136, 377-378, 1993
 A:Title: Characterization of the transcripts encoding two isoforms of human microtubule-associated protein 2, splice form MAP-2c - human
 A:Reference number: I53693; MUID:94124038; PMID:8294038
 A:Accession: I67793
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: mRNA
 A:Residues: 1-472 <ALB>
 A:Cross-references: GB:LI2563; NID:g1850614; PIDN:AAB48097.1; PID:g1850615
 C:Genetics:
 A:Gene: GDB:MAP2; MAP2A; MAP2B; MAP2C
 A:Cross-references: GDB:II8836; OMIM:157130
 A:Map position: 2q34-2q34
 A:Introns: 151/3
 C:Superfamily: microtubule-associated protein MAP2b; MAP2/tau repeat homology
 F:314-344/Domain: MAP2/tau repeat homology <MT1>
 F:376-407/Domain: MAP2/tau repeat homology <MT3>
 Query Match 19.0%; Score 752; DB 2; Length 472;
 Best Local Similarity 38.1%; Pred. No. 1.4e-27;
 Matches 198; Conservative 68; Mismatches 132; Indels 122; Gaps 17;
 QY 297 APL-EFTTFHVEIIPNVQKEQAHSSEHLGRA--FP-----GAPCE-GPEARGPSLGEDT 346
 Db 17 APLTEASAHSH--PPELTQDGGAGEGLVRANGFPPYDEGAFGEHGSQCTVSNTKENG 74
 QY 347 KEADLPSPSKQPAAPAPKGVSRVLPOLKARMSKSKDGTGSDDKKAKTSTRSSAKTLKN 406


```
Db 75 INGEL--TSADRETAEEVSARIVQVVTAEAVAVLK-----GQEKEAQHKDQTAAL----- 123
Qy 407 RCLSPKLTTPGSDPLIQSPSPAVCPPEPPSPKHHVSSVTSTGSSG-----KEMK 458
Db 124 -----PLAAETANLPPSPPPSPASEQVTVEEAAGGESALAPSVFKQAK 168
Qy 459 LKGADGKTKIATPRGAAP-----PCQKG-----QANATRIPAKTP--- 493
Db 169 DKVSDGVTKSPKRRSLPRSSILPRRGVSGDRDENESFSLNSSISSSARRTTTSEPIRR 228
Qy 494 -----PAPKTPPSS-----GEPKSGDRSGYSSPGTSGSRRTPSLTPP-----TR 538
Db 229 AGKSETSTPTPGSTAITPGTPSYSSR-----TPGTPGTP-SYPRTPHTPGTPKSAIILVP 283
Qy 539 EPKKAIVVTPPKSPSSAKSRLOAPVPMPLKNVSKIGSTENLKHQPGGKVOIINKK 598
Db 284 SEKKVAIIRTPPKSPGLTPKQLRLINQPLDULKNVSKIGSTDNIKYQPGGQVIVTKK 343
Qy 599 LDLSNVQSKCGSKDNIKHVPGGGSQVIYKPVDSLKVTSCGSLGNIHHKPGGGGOVEVKS 658
Db 344 IDLSH-----VTSKCSLKNIRHPGGGRVKIES 372
Qy 659 EKLDFKDRVQSKIGSLDNIHVPGGGNKKIETHKLTFRENAKAKTDHGAEIVYKSPVVS 718
Db 373 VKLDFEKAQAQVGSLDNAHHVPGGNNVKIDSKLNFREHAKARVDHGAEIITQSPGRSS 432
Qy 719 DTSRHLNVSTGSDNVDSPLATLADEVSAKQGL 758
Db 433 VASPRLSNVSSGSINLLESPLATLAEDVTAALAKQGL 472
```

Search completed: January 28, 2003, 14:20:12
Job time : 29.8765 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:03:18 ; Search time 11.5946 Seconds
(without alignments)
2711.515 Million cell updates/sec

Title: US-09-904-987-5
Perfect score: 3965
Sequence: 1 MAEPQFEFVEMDHAGTYGL.....SPQLATLADEVSLAKQGL 758

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3960	99.9	757	1 TAU_HUMAN	P10636 homo sapien
2	2852	71.9	751	1 TAU_RAT	P19332 rattus norv
3	2782.5	70.2	732	1 TAU_MOUSE	P10637 mus musculu
4	2032.5	51.3	458	1 TAU_MACMU	P57786 macaca mula
5	1840	46.4	447	1 TAU_BOVIN	P29172 bos taurus
6	1744.5	44.0	382	1 TAU_PAPHA	Q9MYX8 papio hamad
7	1711.5	43.2	402	1 TAU_CAPIH	O02828 capra hircu
8	937	23.6	1861	1 MAP2_RAT	P15146 rattus norv
9	815	20.6	1827	1 MAP2_HUMAN	P11137 homo sapien
10	807	20.4	1828	1 MAP2_MOUSE	P20357 mus musculu
11	624	15.7	1152	1 MAP4_HUMAN	P27816 homo sapien
12	621.5	15.7	1125	1 MAP4_MOUSE	P27546 mus musculu
13	545	13.7	1072	1 MAP4_BOVIN	P36225 bos taurus
14	297	7.5	1411	1 TCOF_HUMAN	Q13428 homo sapien
15	294	7.4	1464	1 CAL3_MOUSE	P08121 mus musculu
16	290	7.3	1049	1 CAL3_BOVIN	P04258 bos taurus
17	287.5	7.3	1806	1 CAL1_HUMAN	P12107 homo sapien
18	283.5	7.2	779	1 CAL1_BOVIN	P02453 bos taurus
19	283.5	7.2	1466	1 CAL1_HUMAN	P02461 homo sapien
20	283	7.1	1185	1 DRPL_HUMAN	P54259 homo sapien
21	282	7.1	1262	1 CAL3_CHICK	P12105 gallus gall
22	279.5	7.0	1685	1 CAS4_HUMAN	P29400 homo sapien
23	278.5	7.0	1453	1 CAL1_CHICK	P02457 gallus gall
24	278.5	7.0	1804	1 CALB_MOUSE	O61245 mus musculu
25	277.5	7.0	1763	1 CA24_ASCSU	P27393 ascaris suu
26	275	6.9	911	1 CALB_BOVIN	Q28083 bos taurus
27	273.5	6.9	1210	1 AF4_HUMAN	P51825 homo sapien
28	273	6.9	1838	1 CAL5_HUMAN	P20908 homo sapien
29	271	6.8	1459	1 CAL2_MOUSE	P28481 mus musculu
30	268.5	6.8	1736	1 CA2B_HUMAN	P13942 homo sapien
31	268	6.8	1183	1 DRPL_RAT	P54258 rattus norv
32	268	6.8	2944	1 CAL7_HUMAN	Q02388 homo sapien
33	267.5	6.7	660	1 YHLL_EBV	P03181 Epstein-bar

ALIGNMENTS

RESULT 1

TAU_HUMAN	6.7	1650	1	CA2B_MOUSE	Q64739
34	267.5	6.7	1464	1	P02452
35	265	6.7	1464	1	homo sapien
36	265	6.7	1669	1	P02463
37	264	6.7	1453	1	mus musculu
38	264	6.7	1453	1	P11087
39	263.5	6.6	1460	1	mus musculu
40	263.5	6.6	1603	1	Q9XSI7
41	263.5	6.6	1581	1	canis famil
42	261.5	6.6	674	1	Q07092
43	261.5	6.6	671	1	homo sapien
44	261	6.6	747	1	Q15648
45	261	6.6	1418	1	h peroxisom
					P33206
					bos taurus
					P02454
					rattus norv
					Q03692
					bos taurus
					Q03692
					homo sapien
					P02458
					homo sapien

STANDARD; PRT: 757 AA.
Q9UJH0;
01-JUL-1989 (Rel. 11, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
15-JUN-2002 (Rel. 41, Last annotation update)
Microtubule-associated protein tau (Neurofibrillary tangle protein)
(Paired helical filament-tau) (PHF-tau).
MAPT OR MTBT1 OR TAU.
Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A. (ISOFORMS PNS-TAU; TAU-A AND TAU-F).
Andreadis A.;
Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
[2]
SEQUENCE FROM N.A. (ISOFORM TAU-A).
TISSUE=Brain;
MEDLINE=88234557; PubMed=3131773;
Goedert M., Wischik C., Crowther R., Walker J., Klug A.;
"Cloning and sequencing of the cDNA encoding a core protein of the
paired helical filament of Alzheimer disease: identification as the
microtubule-associated protein tau";
proc. Natl. Acad. Sci. U.S.A. 85:4051-4055(1988).
[3]
SEQUENCE FROM N.A. (ISOFORMS TAU-B; TAU-C; TAU-E AND TAU-F).
TISSUE=Brain;
MEDLINE=90380393; PubMed=2484340;
Goedert M., Spillantini M.G., Jakes R., Rutherford D., Crowther R.A.;
"Multiple isoforms of human microtubule-associated protein tau:
sequences and localization in neurofibrillary tangles of Alzheimer's
disease";
Neuron 3:519-526(1989).
[4]
SEQUENCE FROM N.A. (ISOFORM TAU-D).
TISSUE=Brain;
MEDLINE=89251564; PubMed=2498079;
Goedert M., Spillantini M.G., Potier M.C., Ulrich J., Crowther R.A.;
"Cloning and sequencing of the cDNA encoding an isoform of
microtubule-associated protein tau containing four tandem repeats:
differential expression of tau protein mRNAs in human brain";
EMBO J. 8:393-399(1989).
[5]
SEQUENCE FROM N.A. (ISOFORMS TAU-A AND FETAL-TAU).
TISSUE=Fetal brain;
MEDLINE=90180482; PubMed=2516729;
Lee G., Neve R.L., Kosik K.S.;
"The microtubule binding domain of tau protein";
Neuron 2:1615-1624(1989).
[6]
SEQUENCE FROM N.A. (ISOFORM TAU-F), AND ALTERNATIVE SPLICING.
MEDLINE=93041757; PubMed=1420178;

- RA Andreadis A., Brown W.M., Kosik K.S.;
 RT "Structure and novel exons of the human tau gene.";
 RL Biochemistry 31:10626-10633(1992).
 RN [7]
 RP SEQUENCE OF 591-621 FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=89193714; PubMed=2495000;
 RA Mori H., Hamada Y., Kawaguchi M., Honda T., Kondo J., Ihara Y.;
 RT "A distinct form of tau is selectively incorporated into Alzheimer's
 RL paired helical filaments.";
 RL Biochem. Biophys. Res. Commun. 159:1221-1226(1989).
 RN [8]
 RP SEQUENCE OF 1-72; 102-380; 467-496; 507-570; 576-582; 591-606;
 RX 615-633; 638-656; 660-663; 670-699 AND 702-757.
 RA TISSUE=Brain;
 RX MEDLINE=92381012; PubMed=1512244;
 RA Hasegawa M., Morishima-Kawashima M., Takio K., Suzuki M., Titani K.,
 RA Ihara Y.;
 RT "Protein sequence and mass spectrometric analyses of tau in the
 RL Alzheimer's disease brain.";
 RL J. Biol. Chem. 267:17047-17054(1992).
 RN [9]
 RP SEQUENCE OF 576-583; 607-610; 615-627; 638-647 AND 670-685,
 RX PHOSPHORYLATION, AND MUTAGENESIS.
 RA Drewes G., Trinczek B., Illenberger S., Biernat J., Schmitt-Ulms G.,
 RA Meyer H.E., Mandelkow E.-M., Mandelkow E.;
 RT "Microtubule-associated protein/microtubule affinity-regulating kinase
 RL interactions and dynamic instability that regulates tau-microtubule
 RL Alzheimer-specific site serine 262.";
 RL J. Biol. Chem. 270:7679-7688(1995).
 RN [10]
 RP REVIEW.
 RX MEDLINE=91320377; PubMed=1713721;
 RA Goedert M., Crowther R.A., Garner C.C.;
 RT "Molecular characterization of microtubule-associated proteins tau and
 RL MAP2";
 RL Trends Neurosci. 14:193-199(1991).
 RN [11]
 RP SUBCELLULAR LOCATION, AND PHOSPHORYLATION.
 RX MEDLINE=20283597; PubMed=10747907;
 RA Maas T., Eidenmueller J., Brandt R.;
 RT "Interaction of tau with the neural membrane cortex is regulated by
 RL phosphorylation at sites that are modified in paired helical
 RL filaments.";
 RL J. Biol. Chem. 275:15733-15740(2000).
 RN [12]
 RP PHOSPHORYLATION, AND MUTAGENESIS.
 RX MEDLINE=98413833; PubMed=9735171;
 RA Sengupta A., Kabat J., Novak M., Wu Q., Grundke-Iqbal I., Iqbal K.;
 RT "Phosphorylation of tau at both Thr 231 and Ser 262 is required for
 RL maximal inhibition of its binding to microtubules";
 RL Arch. Biochem. Biophys. 357:299-309(1998).
 RN [13]
 RP PHOSPHORYLATION, AND MUTAGENESIS.
 RX MEDLINE=98278959; PubMed=9614189;
 RA Illenberger S., Zheng-Fischer O., Preuss U., Stamer K., Baumann K.,
 RA Trinczek B., Biernat J., Godemann R., Mandelkow E.-M., Mandelkow E.;
 RT "The endogenous and cell cycle-dependent phosphorylation of tau
 RL protein in living cells: implications for Alzheimer's disease.";
 RL Mol. Biol. Cell 9:1495-1512(1998).
 RN [14]
 RP GLYCATION.
 RX MEDLINE=97465580; PubMed=9326300;
 RA Nacharaju P., Ko L., Yen S.H.;
 RT "Characterization of in vitro glycation sites of tau.";
 RL J. Neurochem. 69:1709-1719(1997).
 RN [15]
 RP REVIEW ON VARIANTS.
 RX MEDLINE=20437008; PubMed=10899436;
 RA Goedert M., Spillantini M.G.;
 RT "Tau mutations in frontotemporal dementia FTDP-17 and their relevance
 RN for Alzheimer's disease.";
 RL Biochim. Biophys. Acta 1502:110-121(2000).
 RN [16]
 RP VARIANT FTDP17 M-653, AND VARIANTS N-284; A-288; Y-440 AND P-446.
 RX MEDLINE=98291804; PubMed=9629852;
 RA Poorkaj P., Bird T.D., Wijsman E., Nemens E., Garruto R.M.,
 RA Anderson L., Andreadis A., Wiederholt W.C., Raskind M.,
 RA Schellenberg G.D.;
 RT "Tau is a candidate gene for chromosome 17 frontotemporal dementia.";
 RL Ann. Neurol. 43:815-825(1998).
 RN [17]
 RP ERRATUM.
 RA Poorkaj P., Bird T.D., Wijsman E., Nemens E., Garruto R.M.,
 RA Anderson L., Andreadis A., Wiederholt W.C., Raskind M.,
 RA Schellenberg G.D.;
 RL Ann. Neurol. 44:428-428(1998).
 RN [18]
 RP VARIANT FTDP17 LEU-617.
 RX MEDLINE=98409513; PubMed=9736786;
 RA Dumanchin C., Camuzat A., Campion D., Verpillat P., Hannequin D.,
 RA Dubois B., Saugier-Verber P., Martin C., Penet C., Charbonnier F.,
 RA Agid Y., Frebourg T., Brice A.;
 RT "Segregation of a missense mutation in the microtubule-associated
 RL protein tau gene with familial frontotemporal dementia and
 RL parkinsonism.";
 RL Hum. Mol. Genet. 7:1825-1829(1998).
 RN [19]
 RP VARIANTS FTDP17 VAL-588; LEU-617 AND TRP-722.
 RX MEDLINE=98303385; PubMed=9641683;
 RA Hutton M., Lendon C.L., Rizzu P., Baker M., Froelich S., Houlden H.,
 RA Pickering-Brown S., Chakraverty S., Isaacs A., Grover A., Hackett J.,
 RA Adamson J., Lincoln S., Dickson D., Davies P., Petersen R.C.,
 RA Stevens M., de Graaff E., Wauters E., van Baren J., Hillebrand M.,
 RA Joosse M., Kwon J.M., Nowotny P., Che L.K., Norton J., Morris J.C.,
 RA Read L.A., Trojanowski J.J., Basun H., Lannfelt L., Neystat M., Fahn S.,
 RA Dark F., Tannenberg T., Dodd P.R., Hayward N., Kwok J.B.J.,
 RA Schofield P.R., Andreadis A., Snowden J., Craufurd D., Neary D.,
 RA Owen F., Oostra B.A., Hardy J., Goate A., van Swieten J., Mann D.,
 RA Lynch T., Heutink P.;
 RT "Association of missense and 5'-splice-site mutations in tau with the
 RL inherited dementia FTDP-17.";
 RL Nature 393:702-705(1998).
 RN [20]
 RP VARIANT PPND LYS-595, AND VARIANT FTDP17 LEU-617.
 RX MEDLINE=99007274; PubMed=9789048;
 RA Clark L.N., Poorkaj P., Wszolek Z., Geschwind D.H., Nasreddine Z.S.,
 RA Miller B., Li D., Payami H., Awert F., Markopoulou K., Andreadis A.,
 RA D'Souza I., Lee V.M.-Y., Reed L., Trojanowski J.Q., Zhukareva V.,
 RA Bird T., Schellenberg G., Wilhelmsen K.C.;
 RT "Pathogenic implications of mutations in the tau gene in
 RL pallido-ponto-nigral degeneration and related neurodegenerative
 RL disorders linked to chromosome 17";
 RL Proc. Natl. Acad. Sci. U.S.A. 95:13103-13107(1998).
 RN [21]
 RP VARIANTS FTDP17 VAL-588; LYS-596 DEL; LEU-617 AND TRP-722.
 RX MEDLINE=99138654; PubMed=9973279;
 RA Rizzu P., Van Swieten J.C., Joosse M., Hasegawa M., Stevens M.,
 RA Tiben A., Niermeijer M.F., Hillebrand M., Ravid R., Oostra B.A.,
 RA Goedert M., van Duljin C.M., Heutink P.;
 RT "High prevalence of mutations in the microtubule-associated protein
 RL tau in a population study of frontotemporal dementia in the
 RL Netherlands";
 RL Am. J. Hum. Genet. 64:414-421(1999).
 RN [22]
 RP VARIANTS FTDP17 LEU-617; MET-653 AND TRP-722.
 RX MEDLINE=99229757; PubMed=10214944;
 RA Nacharaju P., Lewis J., Easson C., Yen S., Hackett J., Hutton M.,
 RA Yen S.H.;
 RT "Accelerated filament formation from tau protein with specific FTDP-17
 RL missense mutations.";
 RL FEBS Lett. 447:195-199(1999).
 RN [23]
 RP VARIANT FTDP17/CBD SER-617.

Query Match 99.9%; Score 3960; DB 1; Length 757;
Best Local Similarity 100.0%; Pred. No. 2.2e-144;
Matches 757; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AEPQREFEVMEHAGTYGLDRKDDGGYTMHQDEGDTAGLKESPLQTPTDSEEPGS 61
DB 1 AEPQREFEVMEHAGTYGLDRKDDGGYTMHQDEGDTAGLKESPLQTPTDSEEPGS 60

QY 62 ETSDAKSTPTAEDVTAPLVDEGAPGKQAAAPHTPEGTABEAGIGDTPSLEDAAGH 121
DB 61 ETSDAKSTPTAEDVTAPLVDEGAPGKQAAAPHTPEGTABEAGIGDTPSLEDAAGH 120

QY 122 VTQPESSKVVQEGFLREPGLSHQMSGMPGAPLLPEGPREATRQPSGTGPEDETEGG 181
DB 121 VTQPESSKVVQEGFLREPGLSHQMSGMPGAPLLPEGPREATRQPSGTGPEDETEGG 180

QY 182 RHAPELLKHOLLGDLHQEGPPLKAGGKERPGSKKEEVEDRDVDESSPQSPSKASPAQ 241
DB 181 RHAPELLKHOLLGDLHQEGPPLKAGGKERPGSKKEEVEDRDVDESSPQSPSKASPAQ 240

QY 242 DGRPQTAAREATSIPIGFAEGAIPLVDFLSKVSTEIPASEPDGVSVGRAKQDAPLEF 301
DB 241 DGRPQTAAREATSIPIGFAEGAIPLVDFLSKVSTEIPASEPDGVSVGRAKQDAPLEF 300

QY 302 TFHVEITPNVQKEAHSEHILGRAAPGAPGEGPEARGPSLGEDTKAADLPEPSEKQPA 361
DB 301 TFHVEITPNVQKEAHSEHILGRAAPGAPGEGPEARGPSLGEDTKAADLPEPSEKQPA 360

QY 362 APRGKPVSRVPLKARVMVSKDGTGSDKKAKTSTRSSAKTLKRPCLSPKLTPTPGSSD 421
DB 361 APRGKPVSRVPLKARVMVSKDGTGSDKKAKTSTRSSAKTLKRPCLSPKLTPTPGSSD 420

QY 422 PLIQSSPAPVCPPEPPSPKHVSVTSRTGSSGAKEMKLGADGKTATPRGAAPGQKG 481
DB 421 PLIQSSPAPVCPPEPPSPKHVSVTSRTGSSGAKEMKLGADGKTATPRGAAPGQKG 480

QY 482 QANATRIAPKTPAPKTPPSSGPPKSGDRSGYSSPGSGTSGRSRTPSLTPPTPREPK 541
DB 481 QANATRIAPKTPAPKTPPSSGPPKSGDRSGYSSPGSGTSGRSRTPSLTPPTPREPK 540

QY 542 KVAVRTPPKSPSASRLQATVPMPDLKNVSKIGSTENLKHQGGGKVQIINKKDL 601
DB 541 KVAVRTPPKSPSASRLQATVPMPDLKNVSKIGSTENLKHQGGGKVQIINKKDL 600

QY 602 SNVQSKGSDNKHVPGGSGVOIVKPYDLSKVTSCGSLGNHKKPGGGQVEYKSEKL 661
DB 601 SNVQSKGSDNKHVPGGSGVOIVKPYDLSKVTSCGSLGNHKKPGGGQVEYKSEKL 660

QY 662 DFDRVQSKIGSLDNITHVPGGSKKIEHKLTFRENAKAKTDHGAIEIVYKSPVSGDTS 721
DB 661 DFDRVQSKIGSLDNITHVPGGSKKIEHKLTFRENAKAKTDHGAIEIVYKSPVSGDTS 720

QY 722 PRHLSNVSTGSDIMVDSPOLATLADVSASLAKQGL 758
DB 721 PRHLSNVSTGSDIMVDSPOLATLADVSASLAKQGL 757

RESULT 2
TAU_RAT
ID TAU_RAT STANDARD; PRT; 751 AA.
AC P19332; Q63567; Q9QW06; Q63677;
DT 01-NOV-1990 (Rel. 16, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Microtubule-associated protein tau (Neurofibrillary tangle protein)
DE (paired helical filament-tau) (PHF-tau).
GN MAPT OR MTAPO OR TAU.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]

RP SEQUENCE FROM N.A. (ISOFORM TAU-B).
RC TISSUE=Pheochromocytoma;
RX MEDLINE=92179305; PubMed=1542696;
RA Goedert M., Spillantini M.G., Crowther R.A.;
RT "Cloning of a big tau microtubule-associated protein characteristic of
RL the peripheral nervous system.";
RN Proc. Natl. Acad. Sci. U.S.A. 89:1983-1987(1992).
RP SEQUENCE FROM N.A. (ISOFORM TAU-B).
RC TISSUE=Dorsal root ganglion;
RX MEDLINE=94013081; PubMed=8408300;
RA Georgieff I.S., Liem R.K.H., Couchie D., Mavilia C., Nunez J.,
RA Shelanski M.L.;
RT "Expression of high molecular weight tau in the central and peripheral
RL nervous systems.";
RN J. Cell Sci. 105:729-737(1993).
RP SEQUENCE FROM N.A. (ISOFORM TAU-F).
RC STRAIN=Wistar; TISSUE=Brain;
RX MEDLINE=94334997; PubMed=8057376;
RA Sadot E., Marx R., Barg J., Behar L., Ginzburg I.;
RT "Complete sequence of 3'-untranslated region of tau from rat central
RL nervous system. Implications for mRNA heterogeneity.";
RN J. Mol. Biol. 241:325-331(1994).
RP SEQUENCE FROM N.A. (ISOFORMS TAU-E AND TAU-G).
RC TISSUE=Brain;
RX MEDLINE=90180457; PubMed=2560640;
RA Kosik K.S., Orecchio L.D., Bakalis S., Neve R.L.;
RT "Developmentally regulated expression of specific tau sequences.";
RN Neuron 2:1389-1397(1989).
RP SEQUENCE FROM N.A. (ISOFORMS TAU-E AND TAU-C).
RX MEDLINE=89359509; PubMed=2504728;
RA Kanai Y., Takemura R., Oshima T., Mori H., Ihara Y., Yanagisawa M.,
RA Masaki T., Hirokawa N.;
RT "Expression of multiple tau isoforms and microtubule bundle formation
RL in fibroblasts transfected with a single tau cDNA.";
RN J. Cell Biol. 109:1173-1184(1989).
RP SEQUENCE OF 359-460 FROM N.A. (ISOFORM TAU-A), AND SEQUENCE OF 105-112
AND 367-460 FROM N.A. (ISOFORM TAU-D).
RC TISSUE=Spinal cord;
RX MEDLINE=95054048; PubMed=7964751;
RA Mavilia C., Couchie D., Nunez J.;
RT "Diversity of high-molecular-weight tau proteins in different regions
RL of the nervous system.";
RN J. Neurochem. 63:2300-2306(1994).
RP SEQUENCE OF 696-751 FROM N.A. (ISOFORMS TAU-A TO TAU-G), AND SEQUENCE
OF 751-774 FROM N.A. (ISOFORM TAU-H).
RC STRAIN=Sprague-Dawley; TISSUE=Brain;
RX MEDLINE=95182802; PubMed=7877441;
RA Sawa A., Oyama F., Matsushita M., Ihara Y.;
RT "Molecular diversity at the carboxyl terminus of human and rat tau.";
RN Brain Res. Mol. Brain Res. 27:111-117(1994).
CC -1- FUNCTION: PROMOTES MICROTUBULE ASSEMBLY AND STABILITY, AND MIGHT
CC BE INVOLVED IN THE ESTABLISHMENT AND MAINTENANCE OF NEURONAL
CC POLARITY. THE C-TERMINUS BINDS AXONAL MICROTUBULES WHILE THE N-
CC TERMINUS BINDS NEURAL PLASMA MEMBRANE COMPONENTS, SUGGESTING THAT
CC TAU FUNCTIONS AS A LINKER PROTEIN BETWEEN BOTH. AXONAL POLARITY IS
CC PREDETERMINED BY TAU LOCALIZATION (IN THE NEURONAL CELL) IN THE
CC DOMAIN OF THE CELL BODY DEFINED BY THE CENTROSOME. THE SHORT
CC ISOFORMS ALLOW PLASTICITY OF THE CYTOSKELETON WHEREAS THE LONGER
CC ISOFORMS MAY PREFERENTIALLY PLAY A ROLE IN ITS STABILIZATION.
CC -1- SUBCELLULAR LOCATION: MOSTLY FOUND IN THE AXONS OF NEURONS. IN THE
CC CYTOSOL AND IN ASSOCIATION WITH PLASMA MEMBRANE COMPONENTS,
CC -1- ALTERNATIVE PRODUCTS: AT LEAST 8 ISOFORMS; TAU-A/SC1 (SHOWN HERE),
CC TAU-B/BIG-TAU/HW-TAU, TAU-C, TAU-D/SC2, TAU-E, TAU-F, TAU-
CC GFETAL-TAU AND TAU-H; ARE PRODUCED BY ALTERNATIVE SPLICING. THEY
CC DIFFER FROM EACH OTHER BY THE PRESENCE OR ABSENCE OF UP TO 4 OF
CC THE 14 EXONS. TWO DIFFERENT C-TERMINI ARE OBTAINED EITHER BY THE
CC RETENTION OR THE SPLICING OF INTRON 13/14.

QY 297 APLEFTFHEITPNVQEAHSEHLGRAAFGAPGEGPEA--RGPISLGEDTKADLP 354
 Db 274 AAEFTFHEIKASTPEQ-----DEGATVVGPEGEQAKTQGPSVGKGTKEASLQEP 328
 QY 355 SEKQAAAAPRGKPVSRVQPKARWVSKSDGTGDDKAKTSTRSAKTLKNRCLSPKL 414
 Db 329 PGKQPAAGLPGRPVSRVQPKARV--SKDRGTGNDKAKTSTPSCAKAPSHRCLSPTR 386
 QY 415 PTPGSDPLIOPSPAVCPPEPSPKHSVTSRTGSSCAKEMKLGADGKT--KIATPR 472
 Db 387 PTLGSDPLIKPSPAVSPAPATSPKHVSSTVPRNGSPCTKMKLGADGKTGAKIATPR 446
 QY 473 GAAPPGKQGANATRIAPKTPPAPKTPPGSGEPKSGDGRSGYSSPGSPGTPGSRSTRPSL 532
 Db 447 GAASPAQKGTSNATRIAPKTPPSPKTPPGSGEPKSGEYSSPGSPGTPGSRSTRPSL 506
 QY 533 PTPPTREPKKVAVVTRPPKSPSAKSLQATAPVMPDLKNVSKIGSTENLKHQPGGKV 592
 Db 507 PTPPTREPKKVAVVTRPPKSPSAKSLQATAPVMPDLKNVSKIGSTENLKHQPGGKV 566
 QY 593 QIINKKLDLSNVQSKGSKDNKIKHVPGGGQVQIVYKPDLSKVTSCGSLGNHHPGGG 652
 Db 567 QIINKKLDLSNVQSKGSKDNKIKHVPGGGQVQIVYKPDLSKVTSCGSLGNHHPGGG 626
 QY 653 QVEVSEKLDKDRVQSKIGSLDNITHVPGGNGKTIETHKLTFRENAKAKTDHGAIVYK 712
 Db 627 QVEVSEKLDKDRVQSKIGSLDNITHVPGGNGKTIETHKLTFRENAKAKTDHGAIVYK 686
 QY 713 SPVVGDTSPRHLNVSSTGSDIMVDSPLATLADEVSASLAKQGL 758
 Db 687 SPVVGDTSPRHLNVSSTGSDIMVDSPLATLADEVSASLAKQGL 732

RESULT 4
 ID TAU_MACMU STANDARD; PRT: 458 AA.
 AC P57786;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Microtubule-associated protein tau (Neurofibrillary tangle protein)
 DE (Paired helical filament-tau) (PHF-tau).
 GN MAPT OR TAU.
 OS Macaca mulatta (Rhesus macaque).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Macaca.
 OC NCBI_TaxID=9544;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS A; B; C AND D).
 RC TISSUE=Brain;
 RX MEDLINE=97012131; PubMed=8858947;
 RA Nelson P.T., Stefansson K., Gulcher J., Saper C.B.;
 RT "Molecular evolution of tau protein: implications for Alzheimer's
 disease";
 RL J. Neurochem. 67:1622-1632(1996).
 CC -1- FUNCTION: PROMOTES MICROTUBULE ASSEMBLY AND STABILITY. AND MIGHT
 BE INVOLVED IN THE ESTABLISHMENT AND MAINTENANCE OF NEURONAL
 POLARITY. THE C-TERMINUS BINDS AXONAL MICROTUBULES WHILE THE N-
 TERMINUS BINDS NEURAL PLASMA MEMBRANE COMPONENTS, SUGGESTING THAT
 TAU FUNCTIONS AS A LINKER PROTEIN BETWEEN BOTH. AXONAL POLARITY IS
 PREDETERMINED BY TAU LOCALIZATION (IN THE NEURONAL CELL) IN THE
 DOMAIN OF THE CELL BODY DEFINED BY THE CENTROSOME. THE SHORT
 ISOFORMS ALLOW PLASTICITY OF THE CYTOSKELETON WHEREAS THE LONGER
 ISOFORMS MAY PREFERENTIALLY PLAY A ROLE IN ITS STABILIZATION.
 CC -1- SUBCELLULAR LOCATION: MOSTLY FOUND IN THE AXONS OF NEURONS, IN THE
 CYTOSOL AND IN ASSOCIATION WITH PLASMA MEMBRANE COMPONENTS.
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST 4 ISOFORMS; TAU A (SHOWN HERE),
 TAU-B, TAU-C AND TAU-D; ARE PRODUCED BY ALTERNATIVE SPLICING. THEY
 DIFFER FROM EACH OTHER BY THE PRESENCE OR ABSENCE OF UP TO 4
 EXONS. ONE OF THESE OPTIONAL EXONS CONTAINS THE ADDITIONAL TAU/MAP
 REPEAT. THE SEQUENCE SHOWN HERE IS THAT OF THE COMPLETE ISOFORM

CC NOT YET CHARACTERIZED.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN NEURONS.
 CC -1- DOMAIN: THE TAU/MAP REPEAT BINDS TO TUBULIN. TYPE I ISOFORMS
 CC CONTAIN 3 REPEATS WHILE TYPE II ISOFORMS CONTAIN 4 REPEATS.
 CC -1- PTM: PHOSPHORYLATION AT VARIOUS SERINE AND THREONINE RESIDUES IN
 CC S-P OR T-P MOTIFS BY PROLINE-DIRECTED PROTEIN KINASES (PDPK: CDC2,
 CC CDK5, GSK3, MARK) (A FEW SITES PER PROTEIN IN INTERPHASE, MORE IN
 CC MITOSIS), AND AT SERINE RESIDUES IN K-X-G-S MOTIFS BY
 CC MAP/MICROTUBULE AFFINITY-REGULATING KINASE (MARK) (BY SIMILARITY).
 CC -1- SIMILARITY: CONTAINS 4 TAU/MAP REPEATS.
 DR InterPro: IPR002955; Tau.protein.
 DR InterPro: IPR001084; Tubulin.Tau.
 DR Pfam: PF00418; tubulin-binding; 4.
 DR PRINTS: PRO1261; TAUPROTEIN.
 DR PROSITE: PS00229; TAU.MAP; 4.
 KW Microtubules; Cytoskeleton; Repeat; Alternative splicing; Acetylation;
 KW Phosphorylation.
 FT INIT_MET 0 0 BY SIMILARITY.
 FT REPEAT 261 291 TAU/MAP MOTIF 1.
 FT REPEAT 292 322 TAU/MAP MOTIF 2.
 FT REPEAT 323 353 TAU/MAP MOTIF 3.
 FT REPEAT 354 385 TAU/MAP MOTIF 4.
 FT MOD_RES 1 1 ACETYLATION (BY SIMILARITY).
 FT DISULFID 308 339 BY SIMILARITY.
 FT VARSPLIC 44 101 MISSING (IN ISOFORM TAU-B AND ISOFORM
 TAU-C).
 FT VARSPLIC 185 202 MISSING (IN ISOFORM TAU-C AND ISOFORM
 TAU-D).
 FT VARSPLIC 292 322 MISSING (IN ISOFORM TAU-B).
 SQ SEQUENCE 458 AA; 47841 MW; 68BE63DBA30665A8 CRC64;

Query Match 51.3%; Score 2032.5; DB 1; Length 458;
 Best Local Similarity 55.6%; Pred. No. 2.5e-71;
 Matches 431; Conservative 2; Mismatches 7; Indels 335; Gaps 3;

QY 2 AEPQEFVEMDHAGTYGLGDRKQGGYTMHQDEGTDAGLKESPLQTPTDESGEPGS 61
 Db 1 AEPQEFVEMDHAGTYGLGDRKQGGYTMHQDEGTDAGLKESPLQTPTDESGEELGS 60
 QY 62 ETSDAKSTPTAEDVTAPLVDEGAPGKQAAAPHTPEIGTTAEAGIGDTPSLDEAAGH 121
 Db 61 ETSDAKSTPTAEDVTAPLVDERAFGEQAAAPHTPEIGTTAEAGIGDTPSLDEAAGH 120
 QY 122 VTQEPESGVVQEGFLREPGLPGLSHQLMSGMPCAPLLPEGPREATRQPSGTGPDTEGG 181
 Db 121 VTQ----- 123
 QY 182 RHAPELLKHQLLDLHQEGPLPKGAGGKERPGSKEEVEDRDESSPODSPPSKASPAQ 241
 Db 124 ----- 123
 QY 242 DGRPPQTAAREATSIPTPAGGAIPLPVDFLSKYSTETPASEPDGPSVGRAGQDAPLEF 301
 Db 124 ----- 123
 QY 302 TFHVEITPNVQEAHSEHLGRAAFPGAPGEGPEARGPSIGEDTKADLPSEKQPA 361
 Db 124 ----- 123
 QY 362 APRGKPSRVQPKARWVSKSDGTGSDKKAKTSTRSAKTLKNRCLSPKLTTPGSSD 421
 Db 124 -----ARMVSKSDGTGSDDKKA----- 141
 QY 422 PLIQSPSPAVCPPEPSPKHSVTSRTGSSCAKEMKLGADGKTATPRGAAPQKQ 481
 Db 142 -----KGADGKTATPRGAAPQKQ 163
 QY 482 QANATRIAPKTPPAPKTPPSS-----GPPKSGDRSGYSSPGSPGTP 523
 Db 164 QANATRIAPKTPPAPKTPPSSATKQVQRKPPAPETSERGEPKSGDRSGYSSPGSPGTP 223
 QY 524 GSRSTRPSLPTTPREPKKVAVVTRPPKSPSAKSLQATAPVMPDLKNVSKIGSTENL 583
 Db 524 GSRSTRPSLPTTPREPKKVAVVTRPPKSPSAKSLQATAPVMPDLKNVSKIGSTENL 583

Db 224 GSRSTPPLTPPAREPKVAVVTRTPKSPSSAKSLQATAPVMPDLKNVSKSIGSTENL 283
 QY 584 KHQPGGKQVQIINKLDLNSVQSKGSKDNKHVPGGSGVQIYKVPDLSKVTSGSGLG 643
 Db 284 KHQPGGKQVQIINKLDLNSVQSKGSKDNKHVPGGSGVQIYKVPDLSKVTSGSGLG 343
 QY 644 NIHKPGGGGVVEKSEKLFKRVQSKIGSLDNITHVPGGGNKKIETHKLTFRENAKAKT 703
 Db 344 NIHKPGGGGVVEKSEKLFKRVQSKIGSLDNITHVPGGGNKKIETHKLTFRENAKAKT 403
 QY 704 DHGAEIVKPVVSGDTPSRHLSNVSGTSDIDWDSQPLATLADEVASLAKQGL 758
 Db 404 DHGAEIVKPVVSGDTPSRHLSNVSGTSDIDWDSQPLATLADEVASLAKQGL 458

RESULT 5
 TAU_BOVIN
 ID TAU_BOVIN STANDARD: PRT: 447 AA.
 AC P29172; P29173; Q28185; Q28186; Q28187; Q28188; Q28189; Q28190;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE Microtubule-associated protein tau (Neurofibrillary tangle protein)
 DE (Paired helical filament-tau) (PHF-tau).
 GN MAPT OR TAU.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS TAU-A; TAU-B; TAU-G AND TAU-H).
 RC TISSUE=Brain;
 RX MEDLINE=89261765; PubMed=2498649;
 RA Himmler A., Drechsel D., Kirschner M.W., Martin D.W. Jr.;
 RT "Tau consists of a set of proteins with repeated C-terminal
 RT microtubule-binding domains and variable N-terminal domains.";
 RL Mol. Cell. Biol. 9:1381-1388(1989).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS TAU-A TO TAU-F AND TAU-I TO TAU-T).
 RC TISSUE=Brain;
 RX MEDLINE=89261766; PubMed=2498650;
 RA Himmler A.;
 RT "Structure of the bovine tau gene: alternatively spliced transcripts
 RT generate a protein family.";
 RL Mol. Cell. Biol. 9:1389-1396(1989).
 RN [3]
 RP GLYCOSYLATION
 RX MEDLINE=97067111; PubMed=8910513;
 RA Arnold C.S., Johnson G.V.W., Cole R.N., Dong D.L.-Y., Lee M.,
 RA Hart G.W.;
 RT "The microtubule-associated protein tau is extensively modified with
 RT O-linked N-acetylglucosamine.";
 RL J. Biol. Chem. 271:28741-28744(1996).
 CC -|- FUNCTION: PROMOTES MICROTUBULE ASSEMBLY AND STABILITY, AND MIGHT
 CC BE INVOLVED IN THE ESTABLISHMENT AND MAINTENANCE OF NEURONAL
 CC POLARITY. THE C-TERMINUS BINDS AXONAL MICROTUBULES WHILE THE N-
 CC TERMINUS BINDS NEURAL PLASMA MEMBRANE COMPONENTS, SUGGESTING THAT
 CC TAU FUNCTIONS AS A LINKER PROTEIN BETWEEN BOTH. AXONAL POLARITY IS
 CC PREDETERMINED BY TAU LOCALIZATION (IN THE NEURONAL CELL) IN THE
 CC DOMAIN OF THE CELL BODY DEFINED BY THE CENTROSOME. THE SHORT
 CC ISOFORMS ALLOW PLASTICITY OF THE CYTOSKELETON WHEREAS THE LONGER
 CC ISOFORMS MAY PREFERENTIALLY PLAY A ROLE IN ITS STABILIZATION.
 CC -|- SUBCELLULAR LOCATION: MOSTLY FOUND IN THE AXONS OF NEURONS. IN THE
 CC CYTOSOL AND IN ASSOCIATION WITH PLASMA MEMBRANE COMPONENTS.
 CC -|- ALTERNATIVE PRODUCTS: AT LEAST 20 ISOFORMS; TAU-A/PBT43112 (SHOWN
 CC HERE), TAU-B/PBT43-12, TAU-C, TAU-D, TAU-E, TAU-F, TAU-G/PBT4,
 CC TAU-H/PBT7, TAU-I, TAU-J, TAU-K, TAU-L, TAU-M, TAU-N, TAU-O, TAU-
 CC P, TAU-Q, TAU-R, TAU-S AND TAU-T; ARE PRODUCED BY ALTERNATIVE
 CC SPLICING. THEY DIFFER FROM EACH OTHER BY THE PRESENCE OR ABSENCE
 CC OF UP TO 6 OF THE 14 EXONS. ONE OF THESE OPTIONAL EXONS CONTAINS
 CC THE ADDITIONAL TAU/MAP REPEAT. TAU-A CDNA HAS BEEN CONSTRUCTED
 CC FROM TWO OVERLAPPING CNAPS BY THE AUTHORS OF REF.1. TAU-G AND TAU-

H SEQUENCES BEGIN WITH EXON 6 OR A PART OF IT (EXON 6 IS MISSING
 IN ISOFORMS THAT BEGIN WITH EXON 1). 3 DIFFERENT C-TERMINI ARE
 OBTAINED EITHER BY THE RETENTION OR THE SPLICING OF INTRON 13/14
 (2 DIFFERENT 5' SPLICE DONORS).
 -|- TISSUE SPECIFICITY: EXPRESSED IN NEURONS.
 -|- INDUCTION: DURING NEURITE OUTGROWTH.
 -|- DOMAIN: THE TAU/MAP REPEAT BINDS TO TUBULIN. TYPE I ISOFORMS
 CONTAIN 3 REPEATS WHILE TYPE II ISOFORMS CONTAIN 4 REPEATS.
 -|- PTM: PHOSPHORYLATION AT VARIOUS SERINE AND THREONINE RESIDUES IN
 S-P OR T-P MOTIFS BY PROLINE-DIRECTED PROTEIN KINASES (PDPK: CDC2,
 CK5, GSK3, MAPK) (A FEW SITES PER PROTEIN IN INTERPHASE, MORE IN
 MITOSIS). AND AT SERINE RESIDUES IN K-X-G-S MOTIFS BY
 MAP/MICROTUBULE AFFINITY-REGULATING KINASE (MARK) (BY SIMILARITY).
 -|- PTM: O-LINKED N-ACETYLGLUCOSAMINATION AT MORE THAN 4 SITES PER
 PROTEIN. SITE-SPECIFIC OR STOICHIOMETRIC CHANGES IN GLYCOSYLATION
 MAY MODULATE TAU FUNCTION AND ALSO PLAY A ROLE IN PHF'S FORMATION.
 -|- SIMILARITY: CONTAINS 4 TAU/MAP REPEATS.

 THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 EMBL; L34953; AAA51609.1; -
 EMBL; L34940; AAA51609.1; JOINED.
 EMBL; L34941; AAA51609.1; JOINED.
 EMBL; L34942; AAA51609.1; JOINED.
 EMBL; L34943; AAA51609.1; JOINED.
 EMBL; L34944; AAA51609.1; JOINED.
 EMBL; L34946; AAA51609.1; JOINED.
 EMBL; L34947; AAA51609.1; JOINED.
 EMBL; L34948; AAA51609.1; JOINED.
 EMBL; L34949; AAA51609.1; JOINED.
 EMBL; L34950; AAA51609.1; JOINED.
 EMBL; L34951; AAA51609.1; JOINED.
 EMBL; L34952; AAA51609.1; JOINED.
 EMBL; M26157; AAA30770.1; -
 EMBL; M26158; AAA30771.1; -
 EMBL; M26178; AAA51601.1; ALT_SEQ.
 EMBL; L34940; AAA51601.1; JOINED.
 EMBL; L34941; AAA51601.1; JOINED.
 EMBL; L34942; AAA51601.1; JOINED.
 EMBL; L34943; AAA51601.1; JOINED.
 EMBL; L34944; AAA51601.1; JOINED.
 EMBL; L34946; AAA51601.1; JOINED.
 EMBL; L34947; AAA51601.1; JOINED.
 EMBL; L34948; AAA51601.1; JOINED.
 EMBL; L34949; AAA51601.1; JOINED.
 EMBL; L34950; AAA51601.1; JOINED.
 EMBL; L34951; AAA51601.1; JOINED.
 EMBL; M26178; AAA51602.1; ALT_SEQ.
 EMBL; L34940; AAA51602.1; JOINED.
 EMBL; L34941; AAA51602.1; JOINED.
 EMBL; L34943; AAA51602.1; JOINED.
 EMBL; L34944; AAA51602.1; JOINED.
 EMBL; L34946; AAA51602.1; JOINED.
 EMBL; L34948; AAA51602.1; JOINED.
 EMBL; L34949; AAA51602.1; JOINED.
 EMBL; L34950; AAA51602.1; JOINED.
 EMBL; L34951; AAA51602.1; JOINED.
 EMBL; M26178; AAA51603.1; ALT_SEQ.
 EMBL; L34940; AAA51603.1; JOINED.
 EMBL; L34941; AAA51603.1; JOINED.
 EMBL; L34943; AAA51603.1; JOINED.
 EMBL; L34944; AAA51603.1; JOINED.
 EMBL; L34946; AAA51603.1; JOINED.
 EMBL; L34948; AAA51603.1; JOINED.
 EMBL; L34949; AAA51603.1; JOINED.
 EMBL; L34950; AAA51603.1; JOINED.
 EMBL; L34951; AAA51603.1; JOINED.
 EMBL; M26178; AAA51604.1; ALT_SEQ.

RC TISSUE=Frontal cortex;
RA Wang X.L., Wang J., Schultz C., Hubbard G.B.;
RL Submitted (JUN-2000) to the EMBL/Genbank/DBJ databases.
CC -!- FUNCTION: PROMOTES MICROTUBULE ASSEMBLY AND MAINTENANCE OF NEURONAL
CC BE INVOLVED IN THE ESTABLISHMENT AND MAINTENANCE OF NEURONAL
CC POLARITY. THE C-TERMINUS BINDS AXONAL MICROTUBULES WHILE THE N-
CC TERMINUS BINDS NEURAL PLASMA MEMBRANE COMPONENTS, SUGGESTING THAT
CC TAU FUNCTIONS AS A LINKER PROTEIN BETWEEN BOTH. AXONAL POLARITY IS
CC PREDETERMINED BY TAU LOCALIZATION (IN THE NEURONAL CELL) IN THE
CC DOMAIN OF THE CELL BODY DEFINED BY THE CENTROSOME (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: MOSTLY FOUND IN THE AXONS OF NEURONS, IN THE
CC CYTOSOL AND IN ASSOCIATION WITH PLASMA MEMBRANE COMPONENTS (BY
CC SIMILARITY).
CC -!- TISSUE SPECIFICITY: EXPRESSED IN NEURONS.
CC -!- DOMAIN: THE TAU/MAP REPEAT BINDS TO TUBULIN.
CC -!- PTM: PHOSPHORYLATION AT VARIOUS SERINE AND THREONINE RESIDUES IN
CC S-P OR T-P MOTIFS BY PROLINE-DIRECTED PROTEIN KINASES (PDPK: CDC2,
CC CDK5, GSK3, MAPK) (A FEW SITES PER PROTEIN IN INTERPHASE, MORE IN
CC MITOSIS), AND AT SERINE RESIDUES IN K-X-G-S MOTIFS BY
CC MAP/MICROTUBULE AFFINITY-REGULATING KINASE (MARK) (BY SIMILARITY).
CC -!- SIMILARITY: CONTAINS 4 TAU/MAP REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AF281310; AAF97596.1; -;
DR InterPro: IPR002955; Tau_protein.
DR InterPro: IPR001084; Tubulin_Tau.
DR Pfam: PF00418; tubulin-binding; 4.
DR PRINTS: PR01261; TAUPROTEIN.
DR PROSITE: PS00229; TAU_MAP; 4.
KW Microtubules; Cytoskeleton; Repeat; Acetylation; Phosphorylation.
FT INIT_MET 0 0 BY SIMILARITY.
FT REPEAT 185 215 TAU/MAP MOTIF 1.
FT REPEAT 216 246 TAU/MAP MOTIF 2.
FT REPEAT 247 277 TAU/MAP MOTIF 3.
FT REPEAT 278 309 TAU/MAP MOTIF 4.
FT MOD_RES 1 1 ACETYLATION (BY SIMILARITY).
FT DISULFID 232 263 BY SIMILARITY.
FT SEQUENCE 382 AA; 39879 MW; D2D15A53AA00E8E7 CRC64;

Query Match 44.0%; Score 1744.5; DB 1; Length 382;
Best Local Similarity 49.9%; Pred. No. 1.9e-60;
Matches 378; Conservative 1; Mismatches 3; Indels 375; Gaps 3;

QY 2 AEPROBEFYMEDHAGTYGLGDRKQGYTMHQDQEGDGTDAGLKESPLQTPTEDGSEEPGS 61
Db 1 AEPROBEFYMEDHAGTYGLGDRKQGYTMHQDQEGDGTDAGLKESPLQTPTEDGSEEPGS 61

QY 62 ETSDAKSTPAEDVTAPLVDEGAPGQAAAPHTPEIGTTAAEAGIGDTPSLEDAAGH 121
Db 44 ETSDAKSTPAEDVTAPLVDEGAPGQAAAPHTPEIGTTAAEAGIGDTPSLEDAAGH 121

QY 122 VTQEPESGKVGQEGFLREPPGGLSHQLMSGNGAPLLPEGPREATRPGSTGPDTEGG 181
Db 63 VTQEPESGKVGQEGFLREPPGGLSHQLMSGNGAPLLPEGPREATRPGSTGPDTEGG 181

QY 182 RHAPELLKQLGLDLHQEGPPLKAGGKERPGSKEEVEDRDVDESSPDSPSKASPAQ 241
Db 66 RHAPELLKQLGLDLHQEGPPLKAGGKERPGSKEEVEDRDVDESSPDSPSKASPAQ 241

QY 242 DGRPPQTAREATSTPGPPAEGATPLPVDFLSKVSTPEASEPDGSPVGRAGQADPLEF 301
Db 66 DGRPPQTAREATSTPGPPAEGATPLPVDFLSKVSTPEASEPDGSPVGRAGQADPLEF 301

QY 302 TFHVEITPNVQEAHSEHLGRAAFPAGGEPGAPGSLGDTKEADLPPESEKQPA 361
Db 66 TFHVEITPNVQEAHSEHLGRAAFPAGGEPGAPGSLGDTKEADLPPESEKQPA 361

QY 362 APRGKPVSRVLPOLKARMVSKSDGTGSDDKAKKATSTRSSAKTLKRNPCLSPLKPTPGSSD 421
Db 66 APRGKPVSRVLPOLKARMVSKSDGTGSDDKAKKATSTRSSAKTLKRNPCLSPLKPTPGSSD 421

QY 422 PLIQSSPAVCPPEPPSSPKHVSSVTSRTSGSSGAKEMKLGADGKTKIATPRGAAPPGKG 481
Db 84 PLIQSSPAVCPPEPPSSPKHVSSVTSRTSGSSGAKEMKLGADGKTKIATPRGAAPPGKG 481

QY 482 QANATRIPAKTPPAPKTPPSGEPKSGDRSGYSPGSPGTPGSRSTRPSLPTPTREP 541
Db 106 QANATRIPAKTPPAPKTPPSGEPKSGDRSGYSPGSPGTPGSRSTRPSLPTPTREP 541

QY 542 KVAVVTRTPPKSPSSAKSLQATAPVMPDLKNVSKSGISTENLKHQPGGKGVIIINKL 601
Db 166 KVAVVTRTPPKSPSSAKSLQATAPVMPDLKNVSKSGISTENLKHQPGGKGVIIINKL 601

QY 602 SNVQSKGSKDNKIHVPGGSGVQIVYKPDLSKVTSCGSLGNTHHKPGGQGVKSEKL 661
Db 226 SNVQSKGSKDNKIHVPGGSGVQIVYKPDLSKVTSCGSLGNTHHKPGGQGVKSEKL 661

QY 662 DFKDRVQSKIGSLDNITHVPGGSGNKKIETHKLTRENAKAKTDHGAETIVYKSPVVGDT 721
Db 286 DFKDRVQSKIGSLDNITHVPGGSGNKKIETHKLTRENAKAKTDHGAETIVYKSPVVGDT 721

QY 722 PRHLSNVSSGTSIDWDSPQIATLADEVASLAKOGL 758
Db 346 PRHLSNVSSGTSIDWDSPQIATLADEVASLAKOGL 758

RESULT 7
TAU_CAPHI
ID TAU_CAPHI STANDARD; PRT; 402 AA.
AC 002828;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Microtubule-associated protein tau (Neurofibrillary tangle protein)
DE (Paired helical filament-tau) (PHF-tau).
GN MAPT OR TAU.
OS Capra hircus (Goat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Capra.
OX NCBI_TaxID=9925;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS A AND B).
RC TISSUE=Brain cortex;
RX MEDLINE=97012131; PubMed=8858947;
RA Nelson P.T., Stefansson K., Gulcher J., Saper C.B.;
RT "Molecular evolution of tau protein: implications for Alzheimer's
RT disease";
RL J. Neurochem. 67:1622-1632(1996).
CC -!- FUNCTION: PROMOTES MICROTUBULE ASSEMBLY AND STABILITY, AND MIGHT
CC BE INVOLVED IN THE ESTABLISHMENT AND MAINTENANCE OF NEURONAL
CC POLARITY. THE C-TERMINUS BINDS AXONAL MICROTUBULES WHILE THE N-
CC TERMINUS BINDS NEURAL PLASMA MEMBRANE COMPONENTS, SUGGESTING THAT
CC TAU FUNCTIONS AS A LINKER PROTEIN BETWEEN BOTH. AXONAL POLARITY IS
CC PREDETERMINED BY TAU LOCALIZATION (IN THE NEURONAL CELL) IN THE
CC DOMAIN OF THE CELL BODY DEFINED BY THE CENTROSOME. THE SHORT
CC ISOFORMS ALLOW PLASTICITY OF THE CYTOSKELETON WHEREAS THE LONGER
CC ISOFORMS MAY PREFERENTIALLY PLAY A ROLE IN ITS STABILIZATION.
CC -!- SUBCELLULAR LOCATION: MOSTLY FOUND IN THE AXONS OF NEURONS, IN THE
CC CYTOSOL AND IN ASSOCIATION WITH PLASMA MEMBRANE COMPONENTS.
CC -!- ALTERNATIVE PRODUCTS: AT LEAST 2 ISOFORMS; TAU-A (SHOWN HERE) AND
CC TAU-B; ARE PRODUCED BY ALTERNATIVE SPLICING. THEY DIFFER FROM EACH
CC OTHER BY THE PRESENCE OR ABSENCE OF TWO EXONS, ONE OF THESE
CC OPTIONAL EXONS CONTAINS THE ADDITIONAL TAU/MAP REPEAT.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN NEURONS.
CC -!- DOMAIN: THE TAU/MAP REPEAT BINDS TO TUBULIN. TYPE I ISOFORMS
CC CONTAIN 3 REPEATS WHILE TYPE II ISOFORMS CONTAIN 4 REPEATS.
CC -!- PTM: PHOSPHORYLATION AT VARIOUS SERINE AND THREONINE RESIDUES IN
CC S-P OR T-P MOTIFS BY PROLINE-DIRECTED PROTEIN KINASES (PDPK: CDC2,

CC CDK5, GSK3, MAPK) (A FEW SITES PER PROTEIN IN INTERPHASE, MORE IN
CC MITOSIS) AND AT SERINE RESIDUES IN K-X-G-S MOTIFS BY
CC MAP/MICROTUBULE AFFINITY-REGULATING KINASE (MARK) (BY SIMILARITY).
CC -!- SIMILARITY: CONTAINS 4 TAU/MAP REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: S83347; AAB50785.1; -;
DR InterPro: IPR002955; Tau_protein.
DR InterPro: IPR001084; Tubulin_Tau.
DR Pfam: PF00418; tubulin-binding; 4.
DR PRINTS: PRO1261; TAU-PROTEIN.
DR PROSITE: PS00229; TAU-MAP; 4.
KW Microtubules; Cytoskeleton; Repeat; Alternative splicing; Acetylation;
KW Phosphorylation.
FT INIT_MET 0 0 BY SIMILARITY.
FT REPEAT 205 235 TAU/MAP MOTIF 1.
FT REPEAT 236 266 TAU/MAP MOTIF 2.
FT REPEAT 267 297 TAU/MAP MOTIF 3.
FT REPEAT 298 329 TAU/MAP MOTIF 4.
FT MOD_RES 1 1 ACETYLATION (BY SIMILARITY).
FT DISULFID 252 283 BY SIMILARITY.
FT VARSPPLIC 33 61 MISSING (IN ISOFORM TAU-B).
FT VARSPPLIC 236 266 MISSING (IN ISOFORM TAU-B).
SQ SEQUENCE 402 AA; 41716 MW; 3623B684E9F8AEF CRC64;

Query Match 43.2%; Score 1711.5; DB 1; Length 402;
Best Local Similarity 49.5%; Pred. No. 3.6e-59;
Matches 376; Conservative 8; Mismatches 16; Indels 359; Gaps 6;

QY 2 AEPQEFVEMDHAGTYGLDRKQGGYTMHQDQEGDTAGLKESPLQPTTGGSEEPGS 61
Db 1 AEPQEFVEMDHA-----QGDYTL-QDHEGDMEPGLKESPLQTPADGGSEEPGS 49
QY 62 ETSDAKSTPTAEDVTAPLVDEGAPGQAAQAOPHTPEGTTFEEAGIGTPTSLDEAGH 121
Db 50 ETSDAKSTPTAE-----AEEAGIGTSTNLEDAQAH 80
QY 122 VTQEPESGKVVQEGFLREPGLSHQLMSGMPGAPLLPEGPREATRQPSGTGPEDEGG 181
Db 81 VTQ----- 83
QY 182 RHAPELLKOLLGLHOEGPLKAGGKERPGSKEEVEDRDVDSESSPDSPSKASPAQ 241
Db 84 ----- 83
QY 242 DGRPPQTAAREATISGFFPAEGAIPLVDFLSKVSTEIPASEPDGSPVGRAKQDAPLEF 301
Db 84 ----- 83
QY 302 TFHVEITPNVQKEQAHSEHLGRAFPAGPGEGPEARGPSLGDTEADLPPESEKQPA 361
Db 84 ----- 83
QY 362 APRGKPVSRVQLKARVSKSDGTGDDKKAKTSTRSSAKTLKLRPLSPKLTPTCGSD 421
Db 84 -----ARMVSKGDKGTGDDKKA----- 101
QY 422 PLIQSPFAVCEPPSPFKHVSSTVTSGSSGAKEMKLKADGCK--TKIATPRGAAPQG 479
Db 102 -----KGADGKPGTKTATPRGAAPQG 123
QY 480 KQANATRIPAKTPAPKTPPSGEPKPKSGDRSGYSPGTPGSRSTPSPPTPTRE 539
Db 124 KQANATRIPAKTPPTPTSGTGESCKSGDRSGYSPGTPGSRSTPSPPTPTRE 183
QY 540 PKXVAVVTRTPPKSPSSAKSRLQATVPMPDLKNVKSIGSTENLKHQPGGKQVQIINKKL 599

Db 184 PKXVAVVTRTPPKSPSSAKSRLQAPGMPDLKNVKSIGSTENLKHQPGGKQVQIINKKL 243
QY 600 DLSNVQSKCKGKONIKHVPGGGSGVQIVYKPDLSKVTSCGSLGNTHHKPGGQGVKSE 659
Db 244 DLSNVQSKCKGKONIKHVPGGGSGVQIVYKPDLSKVTSCGSLGNTHHKPGGQGVKSE 303
QY 660 KLDFKDRVQSKIGSLDNIHVPGGGKNIETHKLTFRENAKAKTDHGAEIVYKSPVVS 719
Db 304 KLDFKDRVQSKIGSLDNIHVPGGGKNIETHKLTFRENAKAKTDHGAEIVYKSPVVS 363
QY 720 TSPRHLNVSSTGSDMWDPQLATLADEVSAASLAKOGL 758
Db 364 TSPRHLNVSSTGSDMWDPQLATLADEVSAASLAKOGL 402

RESULT 8
MAP2_RAT
ID MAP2_RAT STANDARD; PRT; 1861 AA.
AC P15146;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Microtubule-associated protein 2 (MAP 2) (MAP2B) (Contains: MAP2C).
GN MAP2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE OF 1-1694 AND 1726-1861 FROM N.A.
RC STRAIN=Wistar; TISSUE=Brain;
RX MEDLINE=90251471; PubMed=2339070;
RA Kindler S., Schwabe B., Schulz B., Garner C.C.;
RT "Complete cDNA sequence encoding rat high and low molecular weight
MAP2.";
RL Nucleic Acids Res. 18:2822-2822(1990).
RN [2]
RP SEQUENCE OF 1-1694 AND 1726-1861 FROM N.A.
RC STRAIN=Wistar; TISSUE=Brain;
RX MEDLINE=91060576; PubMed=2174050;
RA Kindler S., Schulz B., Goedert M., Garner C.C.;
RT "Molecular structure of microtubule-associated protein 2b and 2c from
rat brain.";
RL J. Biol. Chem. 265:19679-19684(1990).
RN [3]
RP SEQUENCE OF 1-151; 1515-1694 AND 1726-1861 FROM N.A.
RX MEDLINE=90221819; PubMed=2326166;
RA Doll T., Papandrikopoulou A., Matus A.;
RT "Nucleotide and amino acid sequences of embryonic rat MAP2c.";
RL Nucleic Acids Res. 18:361-361(1990).
RN [4]
RP DISCUSSION OF SEQUENCE.
RX MEDLINE=89365159; PubMed=2770869;
RA Papandrikopoulou A., Doll T., Tucker R.P., Garner C.C., Matus A.;
RT "Embryonic MAP2 lacks the cross-linking sidearm sequences and
dendritic targeting signal of adult MAP2.";
RL Nature 340:650-652(1989).
RN [5]
RP SEQUENCE OF 1695-1725 FROM N.A.
RX MEDLINE=94110302; PubMed=8282767;
RA Doll T., Meichner M., Riederer B.M., Honegger P., Matus A.;
RT "An isoform of microtubule-associated protein 2 (MAP2) containing
four repeats of the tubulin-binding motif.";
RL J. Cell Sci. 106:633-640(1993).
CC -!- FUNCTION: THE EXACT FUNCTION OF MAP2 IS UNKNOWN BUT MAPS MAY
CC STABILIZE THE MICROTUBULES AGAINST DEPOLYMERIZATION. THEY ALSO
CC SEEM TO HAVE A STIFFENING EFFECT ON MICROTUBULES
CC -!- ALTERNATIVE PRODUCTS: VARIOUS FORMS OF MAP2 ARE PRODUCED BY
CC ALTERNATIVE SPLICING OF THE SAME GENE. MAP2C, THE LOW MOLECULAR
CC FORM OF MAP2, LACKS THE CENTRAL DOMAIN OF MAP2A/B.
CC -!- DEVELOPMENTAL STAGE: MAP2C IS EXPRESSED DURING EMBRYONIC BRAIN
CC DEVELOPMENT AND UNTIL POSTNATAL DAY 10. MAP2B IS EXPRESSED

or send an email to license@isb-sib.ch).

```

CC -----
CC EMBL: U01828; AAA03354.1; -
DR EMBL: U89330; AAB48098.1; -
DR EMBL: U89329; AAB48097.1; -
DR EMBL: M25668; AAA59552.1; -
DR PIR: PL0024; ORHUNT.
DR Genew: HGNC:6839; MAP2.
DR MIM: 157130; -
DR InterPro: IPR001084; Tubulin_Tau.
DR Pfam: PF00418; tubulin-binding; 3.
DR PROSITE: PS00229; TAU_MAP; 2.
KW Microtubules; Repeat; Alternative splicing; Calmodulin-binding.
FT DOMAIN 1447 1467 CALMODULIN-BINDING (POTENTIAL).
FT REPEAT 1661 1691 TAU/MAP MOTIF.
FT REPEAT 1692 1722 TAU/MAP MOTIF.
FT REPEAT 1723 1754 TAU/MAP MOTIF.
FT VARSPLIC 152 1507 MISSING (IN ISOFORM MAP2C).
FT CONFLICT 9 9 A -> G (IN REF. 2).
FT CONFLICT 37 37 A -> G (IN REF. 2).
FT CONFLICT 108 108 A -> G (IN REF. 2).
FT CONFLICT 152 155 MISSING (IN REF. 2).
FT CONFLICT 187 187 S -> K (IN REF. 2).
FT CONFLICT 1655 1655 A -> GL (IN REF. 2).
FT CONFLICT 1736 1736 V -> A (IN REF. 2).
SQ SEQUENCE 1827 AA; 199610 MW; BAC36D0030F5F455 CRC64;

Query Match 20.6%; Score 815; DB 1; Length 1827;
Best Local Similarity 29.4%; Pred. No. 1.6e-24;
Matches 286; Conservative 122; Mismatches 274; Indels 292; Gaps 40;

QY 5 RQFEYMEHAGTYGLGDRKQGGYTMHQDQGGTDAGLKESPLQPTPTDGESEEGSETS 64
DB 926 KDFSV-DKEASHISGDK---SGLSKFEDQKKAN----DRLDTVLEK-SEEHADSK 975
QY 65 KDSFTPTAEDVTAPLVDEGAPKQAAQPHTEIPGTTAEAA--GIGDTPLSEDEAAGHV 122
DB 976 HAKKTEAGD-ETETGLGVTVEQALAK-DLSIPTDASSEKAEGLSVPEI----- 1025
QY 123 TQPESSKVVQEGF-----LREPGPGPLSHOLSMGM-----PGAPLLPEGPRA 166
DB 1026 -AEVPSKKVEQGLDPAVQGLDVKISDFG-----QMASGLNIDRRATELKLEATQDM 1078
QY 167 TROPSTGTGPD-----EGGRHAPELLKHLGLD--LHQEGPLK---GAGGK 209
DB 1079 T--PSSKAQEQADAFMGVESGHMKGKTKYSETEVKQKVPDLVHQEAVDKESYESSE 1136
QY 210 ERPGSKEEVEDRDVESSPDQ-----PP--SKASPAQDGRP----- 245
DB 1137 HESLTWESLKADEGKETSPESSLIQDEIAVKLSVEIPCPVAVSEADLATDERADVQMEF 1196
QY 246 ---PQTAAREATSIQCPAEGALPLPVDFLSKVSTEIPASEPDGPGSVGRAKQDAPLEFT 302
DB 1197 IQGPKESKETPDISTPDSVAEPLHETIVSE-PAEQSEEE---IFAQGEYDKLLER 1251
QY 303 F-----HVEITPNVQK-----BOAHSEEHGRAAPFG 329
DB 1252 SDTLQITDLGVCARGEFVETCPSEHKGVIESVWTIEDDFITVQVTTTDEGESGSHSVRF 1311
QY 330 APEGEPARGPSLGDTKEADLPPESEKQ-----FAAPRGKPVSRVPLKARWVS 380
DB 1312 AALEQPEVERRPSPHDEEEFEVEEAAEAQAPKDGSPAPASPEREVA-LSEYETETVD 1370
QY 381 KSKDGTGSDD-----KKAATSTRSS-AKTLKNRP 408
DB 1371 DYKDETTIDDSINDADSLWVDOTDDRSIMTEQLETIPKEAKAEKARSSLEKHKRKEP 1430
QY 409 C-----LSPKLPPTGSSDPL-----IQSPSPA--VCPEPP 436
DB 1431 FKTRGRISTPERKVAKEPSTVSRDEVRRKKAVYKAEKAKTEVQAHSRPSRKFTLKPA 1490
QY 437 ---SSPKHVSSTVSRGSSGA-----KMKLKGADGKTKIATPRGAAP-----PG 478

```

```

DB 1491 IKYTRPHTLSVVRKRTTAAGGESALAPSVFKQAKVKVSDGVTKSPKRRSLPRPSILPP 1550
QY 479 QKG-----QANATRIPAKTP-----PAKTPPSS-----GEPPKSG 509
DB 1551 RRGVSGDRDENSESLNSSISSARRTRRSEPIRRACKSGTSTPTTGCSTAITGCTPPSY 1610
QY 510 DRSGYSSPGTPTGSRSTRPSLTPP-----TREPKKVAVVRTPPKSPSSAKSRLOTAP 564
DB 1611 SR----TPGTPGTP-SYRPTPTPTGTPKSAIILVPSEKKAIVIRTPPKSPATPK-QLRLIN 1664
QY 565 VPMPLKNNKSKIGSTENLKHQPGGGKVQIINKKLDLSNVQSCGSKDNLIKHHVPGGVSQ 624
DB 1665 QPLPDLKNNVSKIGSTDNIIKYPKGGQVIVTKKIDLSH----- 1703
QY 625 IVYRPVDSLKVTSCGSLGNIHHKPGGGQVEVKSEKIDFKDRVQSKIGSLDNITHVPGGG 684
DB 1704 -----VTSKCGSLKNIRHRPGGGRVKIESVKLDFEKVQAKVGSLDNAHHVPGGG 1753
QY 685 NKIETHKLTFRENAKARTDHGAIEVYKSPVSGDTSRPHLSNVSTGSDMVDSPOLAT 744
DB 1754 NVKIDSQKLNFRHAKARVDHGABIIITQSPGRSSVAPRRLSNVSSSGSINLLESPOLAT 1813
QY 745 LADEVASLAKOGL 758
DB 1814 LAEDVTAALAKOGL 1827

RESULT 10
MAP2_MOUSE
ID MAP2_MOUSE STANDARD; PRT; 1828 AA.
AC P20357;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Microtubule-associated protein 2 (MAP 2).
GN MAP2 OR MAP2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89083571; PubMed=3205744;
RA Wang D., Lewis S.A., Cowan N.J.;
RT "Complete sequence of a cDNA encoding mouse MAP2.";
RL Nucleic Acids Res. 16:11369-11370(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89043973; PubMed=3142041;
RA Lewis S.A., Wang D., Cowan N.J.;
RT "Microtubule-associated protein MAP2 shares a microtubule binding
motif with tau protein.";
RL Science 242:936-939(1988).
CC -!- FUNCTION: THE EXACT FUNCTION OF MAP2 IS UNKNOWN BUT MAPS MAY
CC STABILIZE THE MICROTUBULES AGAINST DEPOLYMERIZATION. THEY ALSO
CC SEEM TO HAVE A STIFFENING EFFECT ON MICROTUBULES.
CC -!- SIMILARITY: CONTAINS 3 TAU/MAP REPEATS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch)
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M21041; AAA39490.1; -
DR PIR: S06467; S06467.
DR PIR: A40115; A40115.
DR MGD: MGI:97175; Mtap2.
DR InterPro: IPR001084; Tubulin_Tau.
DR Pfam: PF00418; tubulin-binding; 3.
DR PROSITE: PS00229; TAU_MAP; 2.
DR

```


QY 43 LKESPLQTPTEDGS-----EE--PGSETSDAKSTPTAEDVTAPLVDEGAPGKAAQPHTE 96


```
Db 317 LPTPEDESSTKDVAPPMEBEIIVGNDDTTPSKKETET-----TLPIKMDLAPPEDVLLTKETE 372
Qy 97 IPEGTTABEAGIGDTPSLEDEAAGHVTPQESGKVVQEGFLREPQPPGLSHQMSGMP-- 154
Db 373 L-----APAKGWSLSEIEEALAKNDVRSAEI-PVAQETVSVSETEVIAVEVLPSPDIT 426
Qy 155 -----GAPLLPEGPREATROPSPGDETEGGRHAPELLKHQLGLDLHGQGPPLKAG- 207
Db 427 TLTKDVTLPLEAERPLVDMTPS-LETEWTLGKETAPTEIN--LG-WAKDMSPLPESEV 482
Qy 208 --GKE-----RPGSKEEVEDRDVDESSPQSPSPK----- 236
Db 483 TLGKDVILPETKVAEFNNVTPLSEEVTSVKDMSPSAETAPLAKNADLHSGTELIVDN 542
Qy 237 -ASPAQD-----GR-----PQTAAREAT 254
Db 543 SNAPASDLALPLETKVAIVPIKDKTQVTEEKPRDSQSLAQMSHQGQSTVPCTASPE-- 600
Qy 255 SIIPGPAEAGIPLPVDFLSKYST-----EIPASEPDGPGSVGRAKGDQADPLEFTFHVETPN 310
Db 601 --PVKAAEQMSTLPIDAPSPLENLEQKETPGSQSPSEPCSGVSRQEAKEA--VGVTGN 654
Qy 311 VQKQAHSEELGRAAFQAGCEGPEARG-----PSLGEDTKADLPPESEKOPA--- 360
Db 655 DITTPPNKE-----PPSPKAKKPLATTQPAKTSTSAKTQTSPLPKQAPPTT 703
Qy 361 -----AAPRGKP-----VSRVPOLKARMYSKSKDGTGSD--DKKAKTS 396
Db 704 SGLNKKPMSLAGSVPAAPHKPAATATARPSTLPARDV-KPKPITEAKVAEKRTSPS 762
Qy 397 TRSSARTLKNRCLPKPLTPGSSDPLIQ--PS--SPA-VCPEPPSS-----PKHVSSV 445
Db 763 KPSSAPALKPGKPTTPTVSKATSPSTLYSTGTPSSRSPATTLPKRPTSIKTEGKPADVKRM 822
Qy 446 TSRTSSGAKEMKLGADGKTKIATPRGAAPPGQKQANATRIPAKTPPAKTPPSSGEP 505
Db 823 TAKSADLSRSKTTSSASSVKRNTTTPGAAPP--AGMTSTRV-----KPMASAPSRSSG 873
Qy 506 PKSGDRSGVSSPGSPGTPGSRSTRPSLTPTPTPREPKVAVVTRTPKPSAKSLQATAPV 565
Db 874 ALSVDKK-----PT-----STRPSSAPRVSLRATT-V 900
Qy 566 PMPDLKNVSKITGSTEENLKHOPGG----- 589
Db 901 SAPDLKSVRSKVGSTENIKHQGGRAKVEKTEAATTAGKEPNNAVTKAAGSIASAKP 960
Qy 590 --GKQIINKKLDLNVQSGKSKDNIKHVPGGSGVQIVYKPVLDLSKVTSCGSLGNTHH 647
Db 961 PAGKVOIVSKKVSYSHIOSKCVSKDNIKHVPGCGNVQIQNKVDISKVSKCGSKANIKH 1020
Qy 648 KPGGQVEVKSEKLFKDRVQSKIGSLDNITHVPGGGNKKIE 689
Db 1021 KPGGGDVKIESQKLFKEKAQKAVGSLDNVGHFPAGGAVKTE 1062

RESULT 13
MAP4_BOVIN
ID MAP4_BOVIN STANDARD; PRT; 1072 AA.
AC P36225.
DT 01-JUN-1994 (Rel. 29, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Microtubule-associated protein 4 (MAP 4) (Microtubule-associated
protein-U) (MAP-U).
GN MAP4.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=90338002; PubMed=2380192;
```

```
RA Aizawa H., Emori Y., Murofushi H., Kawasaki H., Sakai H.,
RA Suzuki K.;
RT "Molecular cloning of a ubiquitously distributed microtubule-
associated protein with Mr 190,000.";
RL J. Biol. Chem. 265:13849-13855(1990).
RN [2]
RX DOMAINS.
RX MEDLINE=91236765; PubMed=2033072;
RA Aizawa H., Emori Y., Mori A., Murofushi H., Sakai H., Suzuki K.;
RT "Functional analyses of the domain structure of microtubule-associated
protein-4 (MAP-U).";
RL J. Biol. Chem. 266:9841-9846(1991).
CC -1- FUNCTION: NON-NEURONAL MICROTUBULE-ASSOCIATED PROTEIN. PROMOTES
MICROTUBULE ASSEMBLY.
CC -1- TISSUE SPECIFICITY: IS DISTRIBUTED UBQUITOUSLY AMONG ALL TISSUES
BUT AMOUNTS ARE LOWER IN CEREBELLUM AND LIVER.
CC -1- PTM: PHOSPHORYLATION OF THE PRO-RICH REGION IN THE C-TERMINUS
NEGATIVELY REGULATES MAP-4 ACTIVITY TO PROMOTE MICROTUBULE
ASSEMBLY.
CC -1- SIMILARITY: CONTAINS 3 TAU/MAP REPEATS.
CC -----
CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D90149; BAA14179.1; -
DR PIR; A37127; A37127.
DR InterPro; IPR001084; Tubulin_Tau.
DR Pfam; PF00418; tubulin-binding; 3.
DR PROSITE; PS00229; TAU_MAP; 3.
KW Microtubules; Repeat; Phosphorylation.
FT DOMAIN 244 530 19 X 14 AA TANDEM REPEATS.
FT REPEAT 244 257 1.
FT REPEAT 258 271 2.
FT REPEAT 272 285 3.
FT REPEAT 286 299 4.
FT REPEAT 300 313 5.
FT REPEAT 314 327 6.
FT REPEAT 328 341 7.
FT REPEAT 342 355 8.
FT REPEAT 384 391 9 (INCOMPLETE).
FT REPEAT 392 405 10.
FT REPEAT 406 417 11.
FT REPEAT 418 431 12.
FT REPEAT 432 445 13.
FT REPEAT 446 460 14.
FT REPEAT 461 474 15.
FT REPEAT 475 488 16.
FT REPEAT 489 502 17.
FT REPEAT 503 516 18.
FT REPEAT 517 530 19.
FT REPEAT 907 937 TAU/MAP MOTIF 1.
FT REPEAT 938 968 TAU/MAP MOTIF 2.
FT REPEAT 969 1000 TAU/MAP MOTIF 3.
SQ SEQUENCE 1072 AA; 111914 MW; E8C17A730989F0D2 CRC64;

Query Match 13.7%; Score 545; DB 1; Length 1072;
Best Local Similarity 26.5%; Pred. No. 1.8e-14;
Matches 199; Conservative 109; Mismatches 273; Indels 170; Gaps 29;

Qy 38 DTDAGLKE-PLQPTEDSGSEEGSETSDAKSTP-----TAEDVTAPLVDEGAPGKQA 89
Db 325 ETDVSLDEDLALSTETETSTVTAQDILLPKFETESIPPTVMDLASAEGTVPDQEMTPKVA 384
Qy 90 AAQPHTEIP-----EGTTAEAGIGTDPSEDEAAGHVTPQESGKVVQEGFLREP 140
Db 385 ASLSEIEAPLDEIDIVSSTEIRSAKEIGL-----SSETEVA----- 419
Qy 141 GPGCLSHQMSGMPGAPLLPEGPPEATROPSPGCTPEDEGGRHAPELLKHQLGLDLHQEC 200
```

```
Db 420 ----LSREM-GLP-----PE--TEAIDKDMAAPPETE-----VIVPVKDMA 454
Qy 201 PPLKGAGGKERGSKEEVEDD---RDVDESSPODS-----PPSKASP 239
Db 455 P-----SPGIETTLAKOVAPRQEELEVLGKDTVSLPATEMALGRNALVPETETVL 505
Qy 240 AOD-GRPPQTAAREA--TSIPGFPAEAGAIPLVPDLSKVSFTEIPASEPDGPGSVGRAGQD 296
Db 506 AKDVAQPETEVNLANNAALAKFSEAEVVPVKDMETAQTQEATSEDSQLKSLQDEQS 565
Qy 297 ABLETFHVEITPNVOKAHESEHLGRAAPGAPGEGPEARGSLGDTKEADLPDPSE 356
Db 566 AVPLMTSEAVVAMQKHSPLTDED---SVLEELBQKKPSSQTSPLPSET--SGVAKPEE 620
Qy 357 KOPAAAPRCKPVSRVQPKARVMSKDGTSDDKKKSTSRSSAKTLKNRPCLSPLPT 416
Db 621 GPPTGSVGNIDITAPNKKELPPSPPEKTKPLATTQPAKTST--SKAQT---QPTSLPKQTA 676
Qy 417 P---GSSDPLIQSPSPAVCPPEPPSPKHHVSSVTSRTSGSGAKEMKLG-ADGK---TKIA 469
Db 677 PTLTGSNK--KPSLASSGVPAAPKRPAAATSRPSTLPKSDTKPKPVAEKIPEKRS 734
Qy 470 TPGAAPGQKQGANATIPAKTP-----PAPKTP---PSSGEPPKSG 509
Db 735 PSKPASAPAVKPGSKSTQAVPKAPATATLASPGSTSRNLSTPLPKRPTAIKTEGPABEK 794
Qy 510 DRSGYSSPGSPCTPGS-----RSRTSLPTPTPREPKVAVVTPPK-----551
Db 795 KWATSPADLRSRKTSTTSVSKKSTTVPGTAPPAGASRAPRATPPRPGTTPVDKPP 854
Qy 552 ---SPSSAKSRL--QTAPVPMPLDNKVKSKIGSTENLKHQPGGKVQIINKKLDS-----602
Db 855 TAAKPTSSAPRLGRVAANASAPDLANVRSKVSTENIKHQPGGGRK-VEKTEAAAPAR 913
Qy 603 ---NVQSKGSKDNKHKVPGGSGVQIVKPVDSLKVTSKGSLGNIHHKPGGGGVEKVS 658
Db 914 KPENAVTKAAGPIGNAQKPPGKVOIQNKVKYDTSKVSCKGSKRANIKHKPGGGVDKIES 973
Qy 659 EKLDPKDRVQSKIGSLDNITHVPGGNGKKIE 689
Db 974 QKLNFEKAQAQVGLDNDVGHLPAGGAVKTE 1004

RESULT 14
TCOF HUMAN
ID TCOF HUMAN STANDARD; PRT; 1411 AA.
AC Q13428; Q99408; Q99860;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Treacle protein (Treacher collins syndrome protein).
GN TCOF1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96154183; PubMed=8563749;
RA Dixon J., Edwards S.J., Gladwin A.J., Dixon M.J., Loftus S.K.,
RA Bonner C.A., Koprivnikar K., Wasmuth J.J.;
RT "Positional cloning of a gene involved in the pathogenesis of
RT Treacher Collins syndrome."
RL Nat. Genet. 12:130-136(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97228900; PubMed=9074926;
RA Dixon J., Edwards S.J., Anderson I., Brass A., Scambler P.J.,
RA Dixon M.J.;
RT "Identification of the complete coding sequence and genomic
RT organization of the Treacher Collins syndrome gene."
RL Genome Res. 7:223-234(1997).
```

```

[3]
RN SEQUENCE FROM N.A.
RX MEDLINE=97250498; PubMed=9096354;
RA Wise C.A., Chiang L.C., Paznekas W.A., Sharma M., Musy M.M.,
RA Ashley J.A., Lovett M., Jabs E.W.;
RT "TCOF1 gene encodes a putative nucleolar phosphoprotein that exhibits
RT mutations in Treacher Collins syndrome throughout its coding
RT region."
RL Proc. Natl. Acad. Sci. U.S.A. 94:3110-3115(1997).
RN [4]
RP VARIANTS L-439; V-810; V-1313 AND G-1355, AND VARIANT TCS R-53.
RX MEDLINE=97195537; PubMed=9042910;
RA Edwards S.J., Gladwin A.J., Dixon M.J.;
RT "The mutational spectrum in Treacher Collins syndrome reveals a
RT predominance of mutations that create a premature-termination
RT codon."
RL Am. J. Hum. Genet. 60:515-524(1997).
CC -1- DISEASE: DEFECTS IN TCOF1 ARE THE CAUSE OF TREACHER COLLINS
CC SYNDROME (TCS). TCS IS A AUTOSOMAL DOMINANT DISORDER OF
CC CRANIOFACIAL DEVELOPMENT THAT OCCURS WITH AN INCIDENCE OF 1/50,000
CC LIVE BIRTHS. THE CLINICAL FEATURES OF TCS ARE BILATERALLY
CC SYMMETRICAL AND INCLUDE: (1) ABNORMALITIES OF THE EXTERNAL EARS,
CC ATRESIA OF THE EXTERNAL EAR CANALS, AND MALFORMATION OF THE MIDDLE
CC EAR OSSICLES, WHICH MAY RESULT IN CONDUCTIVE HEARING LOSS; (2)
CC LATERAL DOWNWARD SLOPING OF PALPEBRAL FISSURES, FREQUENTLY WITH
CC COLOBOMAS OF THE LOWER EYELIDS; (3) HYPOPLASIA OF THE MANDIBLE AND
CC ZYGOMATIC COMPLEX; (4) CLEFT PALATE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U40847; AAC50903.1; -
DR EMBL; U76366; AAC51181.1; -
DR EMBL; U84664; AAC51185.1; -
DR EMBL; U84640; AAC51185.1; JOINED.
DR EMBL; U84641; AAC51185.1; JOINED.
DR EMBL; U84642; AAC51185.1; JOINED.
DR EMBL; U84643; AAC51185.1; JOINED.
DR EMBL; U84644; AAC51185.1; JOINED.
DR EMBL; U84645; AAC51185.1; JOINED.
DR EMBL; U84646; AAC51185.1; JOINED.
DR EMBL; U84647; AAC51185.1; JOINED.
DR EMBL; U84648; AAC51185.1; JOINED.
DR EMBL; U84649; AAC51185.1; JOINED.
DR EMBL; U84650; AAC51185.1; JOINED.
DR EMBL; U84651; AAC51185.1; JOINED.
DR EMBL; U84652; AAC51185.1; JOINED.
DR EMBL; U84653; AAC51185.1; JOINED.
DR EMBL; U84654; AAC51185.1; JOINED.
DR EMBL; U84655; AAC51185.1; JOINED.
DR EMBL; U84656; AAC51185.1; JOINED.
DR EMBL; U84657; AAC51185.1; JOINED.
DR EMBL; U84658; AAC51185.1; JOINED.
DR EMBL; U84659; AAC51185.1; JOINED.
DR EMBL; U84660; AAC51185.1; JOINED.
DR EMBL; U84661; AAC51185.1; JOINED.
DR EMBL; U84662; AAC51185.1; JOINED.
DR EMBL; U79659; AAC51185.1; JOINED.
DR EMBL; U79659; AAB40722.1; -
DR EMBL; U79645; AAB40722.1; JOINED.
DR EMBL; U79646; AAB40722.1; JOINED.
DR EMBL; U79647; AAB40722.1; JOINED.
DR EMBL; U79648; AAB40722.1; JOINED.
DR EMBL; U79649; AAB40722.1; JOINED.
DR EMBL; U79650; AAB40722.1; JOINED.
DR EMBL; U79651; AAB40722.1; JOINED.
DR EMBL; U79652; AAB40722.1; JOINED.
DR EMBL; U79653; AAB40722.1; JOINED.
```

```

DR EMBL; U79654; AAB40722.1; JOINED.
DR EMBL; U79655; AAB40722.1; JOINED.
DR EMBL; U79656; AAB40722.1; JOINED.
DR EMBL; U79657; AAB40722.1; JOINED.
DR EMBL; U79658; AAB40722.1; JOINED.
DR Genew; HGNC:11654; TCOFI.
DR MIM; 606847; -.
DR MIM; 154500; -.
DR InterPro; IPR003993; treacle.
DR Pfam; PF03546; treacle; 3.
DR PRINTS; PRO1503; TREACLE.
KW Disease mutation; Polymorphism.
FT DOMAIN 89 97 POLY-GLU.
FT DOMAIN 204 207 POLY-SER.
FT DOMAIN 616 619 POLY-SER.
FT DOMAIN 919 924 POLY-SER.
FT DOMAIN 1285 1289 POLY-LYS.
FT DOMAIN 1375 1386 POLY-LYS.
FT DOMAIN 1398 1405 POLY-LYS.
FT VARIANT 53 53 W -> R (IN TCS).
FT VARIANT 439 439 /FTID-VAR_005630.
FT VARIANT 810 810 /FTID-VAR_005631.
FT VARIANT 1313 1313 /FTID-VAR_005632.
FT VARIANT 1335 1335 /FTID-VAR_005633.
FT VARIANT 1312 1312 /FTID-VAR_005634.
FT CONFLICT 1312 1312 K -> Q (IN REF. 2).
FT SEQUENCE 1411 AA; 144312 MW; 3880203D985C2699 CRC64;

Query Match 7.5%; Score 297; DB 1; Length 1411;
Best Local Similarity 21.1%; Pred. No. 6.2e-05;
Matches 164; Conservative 92; Mismatches 279; Indels 242; Gaps 33;

QY 1 MAEPFQEFVEMDHAGTYGLGRKDKGGYTMHQDQEGDTPDAGLK-----ESPLQTPTED 54
DB 40 LAQPTLLDIYTHQQTSELGRKR-----KAEDALQAKKTRVSDPIST-SES 87
QY 55 GSEEPGSETSDAKSTPTAEDVTAPLVDEGAP-----GKAAAQPHFTIEPG 100
DB 88 SEEEAEAEATKATPRLASTNSVLGADLPSSMKREKAKAETKAGKTGNSMPHP--ATG 145
QY 101 TTAEEAGIGDTSLEDEAAGH---VTOEPESCKVQVEGFLREPQPPG----- 144
DB 146 KTVANLLSKSPKSNPSANTLTVSETEESVAFGAAGAAKPGMVSAQADSSSEDTS 205
QY 145 -----LSHQLM-----SGMPGAPLLPEGPPE---ATROPSTGTGPDTEGG 181
DB 206 SSETDVEVKASEKILQVRAASAPAKGTPGKATPAPPGKAGAVASQTKAGKPEEDSE 265
QY 182 RH-----APLLKHQLLDLHQBG-----PPLKAG-----GKERP----- 212
DB 266 SESSDSEETPAAKALLOKASGKTSQVGAASAPAKSPESPRGAAPPPGKTGPAVAKA 325
QY 213 GKSEEVDEDRDVEDSSPQSPKSPASQADGRPPPTAAREATSIFFPAEGAIPLPVDFL 272
DB 326 AKGREDSOSSSEESDSEEAQAQKPS--GKAPQVRAASAPAKES-PRKGAAPPAPPRKT 382
QY 273 SKVSTEIPASEPDGPGSVGRAGQDAPLEFTFIVEI-----TPNVQEQAHSEHLGR 324
DB 383 GPAAAQVQGVKQEDSRSSSESDSDREALAAMNAQVKPLGKSPQVRPASTMGWGLPK 442
QY 325 AAFPAGPGE-GPEARGPSLGE-----DTKEADLP---EPEK-----Q 358
DB 443 GAGVPVPGVGPATPSAQVQKWEEDSESSSESDSDGVEPTAVAPAEKSLGNLOAK 502
QY 359 PAAAP--RGKPVSRVP---OLKA-----RMVSKSKDGTGSDDKA----- 393
DB 503 PTSSPAKGPQKAGPVAVQVKAKEKPMDSSESSSDSDSEAEAPAMTAQAQKPAKIP 562
QY 394 -----KTSTRSSAKTLKNRCLSPKLPTGSDPLIQSPSPAVC-----PEPPSPK 440

```

```

Db 563 QTKACPKKTKNTASAKVAPVAVGTPPRKAGTATSP--AGSSPAVAGGTQRPADSSSE 620
QY 441 HVSSVTSRTG-----SSGAKEMKLGADGKTATPRGAAPPQGGQKQANATRIPAKT 492
Db 621 ESDSEEXTGLAVTVGQAKSVGKGLQVKAASVPVK-----GSLGQGTAPVLPCKT 670
QY 493 PP-----APKTPPSSGEPKPSGDRSGVSSP-----GSPG 521
Db 671 GPTVTQVKAKEOEDSESESESESESESESESESESESESESESESESESESESESE 730
QY 522 TPG-----SRRTPSLPTTPPREPKKVVAV-RTPPKSPSSAK-----SRLQATP 564
Db 731 APGVVTTAAQAQKQKSPSKVFPVVRNPONSTVLRGASVPSVGVKAVATAAQATGP 787

RESULT 15
CAL3_MOUSE
ID CAL3_MOUSE STANDARD; PRT; 1464 AA.
AC P08121; Q61429; Q9CRN7;
DT 01-AUG-1988 (Rel. 08, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Collagen alpha 1(III) chain precursor.
GN COL3A1
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6 X DBA; TISSUE=Embryo;
RX MEDLINE=95011609; PubMed=7926795;
RA Tomon D., de Crombrughe B.;
RT "The mouse type-III procollagen-encoding gene: genomic cloning and
RT complete DNA sequence.";
RL Gene 147:161-168(1994).
RN [2]
RP SEQUENCE OF 1-488 FROM N.A.
RX MEDLINE=88167858; PubMed=3443309;
RA Wood L., Theriault N., Vogeli G.;
RT "Complete nucleotide sequence of the N-terminal domains of the murine
RT alpha-1 type-III collagen chain.";
RL Gene 61:225-230(1987).
RN [3]
RP SEQUENCE OF 1-28 FROM N.A.
RX MEDLINE=85131189; PubMed=3972847;
RA Liau G., Mudryj M., de Crombrughe B.;
RT "Identification of the promoter and first exon of the mouse alpha 1
RT (III) collagen gene.";
RL J. Biol. Chem. 260:3773-3777(1985).
RN [4]
RP SEQUENCE OF 810-1464 FROM N.A.
RC STRAIN=Embryonic head;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Kojima Y., Gojibori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;

```

"Functional annotation of a full-length mouse cDNA collection.";

```

RT Nature 409:685-690(2001).
RN [5]
RC SEQUENCE OF 1442-1464 FROM N.A.
RX MEDLINE=91274355; PubMed=2054384;
RA Metaeeranta M., Toman D., de Crombrughe B., Vuorio E.;
RT "Specific hybridization probes for mouse type I, II, III and IX
RL collagen mRNAs";
RL Biochim. Biophys. Acta 1089:241-243(1991).
CC -!- FUNCTION: COLLAGEN TYPE III OCCURS IN MOST SOFT CONNECTIVE TISSUES
CC ALONG WITH TYPE I COLLAGEN.
CC -!- SUBUNIT: TRIMERS OF IDENTICAL ALPHA 1(III) CHAINS. THE CHAINS ARE
CC LINKED TO EACH OTHER BY INTERCHAIN DISULFIDE BONDS. TRIMERS ARE
CC ALSO CROSS-LINKED VIA HYDROXYLYSINES.
CC -!- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING
CC UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
CC O-LINKED GLYCAN CONSISTS OF GLC-GAL DISACCHARIDE (BY SIMILARITY).
CC -!- SIMILARITY: CONTAINS 1 VWFC DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X52046; CAA36279.1; -
DR EMBL: M18933; AAA37338.1; -
DR EMBL: K03037; -; NOT_ANNOTATED_CDS.
DR EMBL: AK019448; BAB31724.1; -
DR EMBL: X57983; CAA41048.1; -
DR PIR: A22287; A22287.
DR PIR: A27353; A27353.
DR PIR: S16373; S16373.
DR MGD: MGI:88453; Col3a1.
DR InterPro: IPR000087; Collagen.
DR InterPro: IPR000885; Fib.collagen_C.
DR InterPro: IPR001007; VWFC.
DR Pfam: PF01391; Collagen; 18.
DR Pfam: PF01410; COLFI; 1.
DR ProDom: PD000007; Collagen; 1.
DR ProDom: PD002078; Fib.collagen_C; 1.
DR SMART: SM00038; COLFI; 1.
DR SMART: SM00214; VWFC; 1.
DR PROSITE: PS01208; VWFC; 1.
KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
KW Glycoprotein; Collagen; Signal.
FT SIGNAL 1 23 BY SIMILARITY.
FT PROPEP 24 154 AMINO-TERMINAL PROPEPTIDE.
FT CHAIN 155 1203 COLLAGEN ALPHA 1(III) CHAIN.
FT PROPEP 1204 1464 CARBOXYL-TERMINAL PROPEPTIDE.
FT DOMAIN 31 90 VWFC.
FT DOMAIN 155 169 NONHELICAL REGION (N-TERMINAL).
FT DOMAIN 170 1195 TRIPLE-HELICAL REGION.
FT DOMAIN 1196 1464 NONHELICAL REGION (C-TERMINAL).
FT CARBOHYD 262 262 O-LINKED (GAL. .) (BY SIMILARITY).
FT MOD_RES 262 262 HYDROXYLATION (BY SIMILARITY).
FT MOD_RES 283 283 HYDROXYLATION (BY SIMILARITY).
FT MOD_RES 859 859 HYDROXYLATION (BY SIMILARITY).
FT MOD_RES 976 976 HYDROXYLATION (BY SIMILARITY).
FT MOD_RES 1093 1093 HYDROXYLATION (BY SIMILARITY).
FT MOD_RES 1105 1105 HYDROXYLATION (BY SIMILARITY).
FT DISULFID 1195 1195 INTERCHAIN (BY SIMILARITY).
FT DISULFID 1196 1196 INTERCHAIN (BY SIMILARITY).
FT SEQUENCE 1464 AA; 138944 MW; 2104EC27A886090B CRC64;
Query Match 7.4%; Score 294; DB 1; Length 1464;
Best Local Similarity 23.2%; Pred. No. 8.3e-05;
Matches 165; Conservative 52; Mismatches 249; Indels 244; Gaps 33;
QY 7 EFWMEDHAGTYGLGRDKRQGGYTHMQDQGGTMDAGLKESPLQTPTEGSE-EPGSETSD 65

```

```

Db 155 QFSDYDVKSGVGMGYPGPAG-----PPGPPGPGSSHPGS----- 192
QY 66 AKSTPTAEDVTAPLVDEGAPCKQA-AAQPHTEIPEGTTAEAGIGDTPTSLDEAAGHVQT 124
Db 193 -----PSGPGYQGGPGEPGQAGPAGPPGPGALG-----PAGPAGK 228
QY 125 EPESGKVVQEGFLREPDPGGLSHOLMSGMPGAPLL----- 159
Db 229 DGESEGRPCRGERCLPGPGGIKGP--AGMGFFPMKCHRGFDGRNGKEGTGAPLKG 286
QY 160 -----PEGPREATQ-----PSGTGPEDETEGRHAPELLKHQLGLDHLQEGPPL 203
Db 287 GLPDGNGAPGPMGRGAPGERGRGLPGAAGARGNDGAR-----GSDGQPGPP- 334
QY 204 KG-----AGGRERPGSKEEVEDRDVDESSPDSPSKASPAQDGRPPQTAARATSI 259
Db 335 -GPPTAGTGGPGSGAKGEV-----GAGSPGNSPGQGEPPGQGHAGAGQPPGP 384
QY 260 PAEGAIPLPVDLFSKVSTEIPASEPDGVSVGRARQDAPLETFTHVEITPNVQKEQAHSE 319
Db 385 PGNNGSPGGKGEVG-----PAGIPGAPGLIGARGPPGAG-----TNGIPCTRGPSG 431
QY 320 E--HLGRAAPGAPGEGPEARGPSL-----GDTKEADLPEP-----SEKQPA----- 360
Db 432 EPGKNGAKGEFGARGERGEAGSGPIPGKGEDGKDGSPGEPGANGLPGAAGERGPGSGFRG 491
QY 361 -AAPRGKPVSRVPOLK-----ARWVSKSDGTGSDDKKA 393
Db 492 PAGNGLPGEKGGPGERGGPGPAGPRGVRGAPGRDGTGPGGIRGMGPGSGGNGDKPG 551
QY 394 KTSTRSSAKTLKNRPCLSPKLTPTGSSDPLIQFSSPAVCPEPPSPKHVSVTSRTGSSG 453
Db 552 PPGSGGES-----GRP-----GPPGSGSPRGQ-----PGVMGFP--GPKGNDGAPCKNGERG 596
QY 454 AK-EMKLGADGKTKIATPRG-----AAPQCKGQANATRIPTAKTPPAK-----T 498
Db 597 GPGGGLPFGPAGKNGETGPGPGPTGTGAGDKGDSG-----PPGPGLOQIGPTGGP 648
QY 499 PPSGGEPPKSGDRSGYSSPGSP-----GTPGSR--RTPSLP-----TPPTREPCKVA 544
Db 649 PGENGKFGEPGPKGEVGAAPGAPGCKGDSGAPGERGPPGTAGIPGARGGAGPPGPGGKGP 708
QY 545 VVTRTPKSPSSAKSRLQTAPEVPMPLKNVKSIG-----STENLKHQPGG 589
Db 709 A--GPPGPPGASGS-----PGLQGMPCERGGGPGSPGPKGEKGEPPG 747

```

Search completed: January 28, 2003, 14:15:17

Job time : 26.5946 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:04:57 ; Search time 38.3037 Seconds
(without alignments)
4077.505 Million cell updates/sec

Title: US-09-904-987-5

Perfect score: 3965

Sequence: 1 MAEPHQEFVEMDHAGTYGL.....SPQLATLADEVASLAKQGL 758

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriaph.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	1537	38.8	372 11 Q91WK4	Q91WK4 mus musculus
2	795	20.1	1825 11 Q64715	Q64715 rattus norv
3	790.5	19.9	1828 11 Q63724	Q63724 rattus ratt
4	715	18.0	323 6 Q28869	Q28869 bos taurus
5	704	17.8	801 6 Q95KD1	Q95KD1 macaca fasc
6	690.5	17.4	198 11 P97749	P97749 rattus sp.
7	576	14.5	381 11 Q64710	Q64710 rattus norv
8	567.5	14.3	414 11 Q9DBX9	Q9DBX9 mus musculus
9	485	12.2	838 5 Q8T5L5	Q8T5L5 anopheles g
10	482.5	12.2	1224 13 Q9YI90	Q9YI90 xenopus lae
11	481	12.1	125 4 Q16296	Q16296 homo sapien
12	471	11.9	700 5 Q9VB15	Q9VB15 drosophila
13	392	9.9	102 11 Q9QVU4	Q9QVU4 rattus sp.
14	379.5	9.6	361 5 Q95UT1	Q95UT1 drosophila
15	377.5	9.5	453 5 O02592	O02592 caenorhabdi
16	371	9.4	928 13 Q98906	Q98906 gallus gall

17	362	9.1	431	5	Q17364	Q17364 caenorhabdi
18	355	9.0	102	11	Q9QVU3	Q9QVU3 rattus sp.
19	350	8.8	339	4	Q8TB56	Q8TB56 homo sapien
20	347	8.8	113	13	Q12967	Q12967 fugu rubrip
21	342.5	8.6	413	5	P90973	P90973 caenorhabdi
22	341	8.6	458	5	Q950G4	Q950G4 caenorhabdi
23	325.5	8.2	436	5	Q17365	Q17365 caenorhabdi
24	309	7.8	2187	11	P70670	P70670 mus musculu
25	299.5	7.6	5120	13	Q9PU36	Q9PU36 gallus gall
26	296	7.5	958	5	Q8SX15	Q8SX15 drosophila
27	294.5	7.4	1422	6	Q95KU4	Q95KU4 canis fami
28	289	7.4	958	4	Q96A52	Q96A52 homo sapien
29	289	7.3	885	5	Q9TXR9	Q9TXR9 caenorhabdi
30	289	7.3	930	5	Q17339	Q17339 caenorhabdi
31	288.5	7.3	847	6	Q95KP7	Q95KP7 canis fami
32	285.5	7.2	990	13	Q91803	Q91803 xenopus lae
33	285	7.2	971	5	Q9XV54	Q9XV54 caenorhabdi
34	285	7.2	1182	4	Q99495	Q99495 homo sapien
35	285	7.2	1835	13	Q91AU4	Q91AU4 gallus gall
36	284.5	7.2	1820	13	Q91907	Q91907 pagrus majo
37	284.5	7.2	2944	11	Q63870	Q63870 mus musculu
38	283.5	7.2	1745	4	Q9NZQ6	Q9NZQ6 homo sapien
39	283	7.1	1838	4	Q15094	Q15094 homo sapien
40	282	7.1	1151	13	O57580	O57580 gallus gall
41	282	7.1	1450	13	Q9YIB4	Q9YIB4 cynops pyrr
42	282	7.1	2703	5	Q9VEG7	Q9VEG7 drosophila
43	281.5	7.1	1445	13	Q93251	Q93251 rana catesb
44	281	7.1	1458	13	Q910B9	Q910B9 oncorhynch
45	281	7.1	1838	11	O88207	O88207 mus musculu

ALIGNMENTS

RESULT 1

Q91WK4 ID Q91WK4 PRELIMINARY; PRT; 372 AA.
AC Q91WK4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Microtubule-associated protein tau.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=EYE, AND RETINA;
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC014748; AAH14748.1; -
DR InterPro; IPR001084; Tubulin_Tau.
DR Pfam; PF00418; tubulin-binding; 4.
DR PROSITE; PS00229; TAU_MAP; UNKNOWN 4.
SQ SEQUENCE 372 AA; 3861 MW; B027745D23BC62A2 CRC64;

Query Match 38.8%; Score 1537; DB 11; Length 372;
Best Local Similarity 45.3%; Pred. No. 3.9e-79;
Matches 344; Conservative 12; Mismatches 14; Indels 390; Gaps 6;
Qy 1 MAEPHQEFVEMDHAGTYGLGDRKQGGYTHQDOEGDTAGLKESPLQPTPEDGSEEPG 60
Db 1 MADPRQEFDTMDHA-----GDYTLQDQEGMDHGLK----- 33
Qy 61 SETSDAKSTPTAEDVTPLVDEGAPGKQAAQAHPTEPGTAAEAGTGTDPSPLEBAAG 120
Db 34 -----AEEAGIGDTPNEDQAAG 51
Qy 121 HVTQEPESGKVVQEGFLREPGLSHQLMSGMPGAPLLPEGPREATRQPSGTGPEDTEG 180
Db 52 HVTQ----- 55

QY 181 GRHAPELLKHQLLDHQBEPPLKAGGKERPGSGKEEVEDDDRDVDESSPDQSPSPKASPA 240
 Db 56 -----TUBULIN-1----- 55
 QY 241 QDGRPQTAAAREATSIPTGPAEGAIPVPDFLSKVSTEIPASEPDGSPVGRAGKQDAPLE 300
 Db 56 -----ARVA----- 59
 QY 301 FTFHVITPNVQEAHSEHLGRAFFAPGAPGEPGARGSLGDETKADLPPESEKQPA 360
 Db 60 ----- 59
 QY 361 AAPRGKPSRVQPKARWVSKSDGTGDDKAKTSTRSSAKTLKNRPCLSPKLPPTGSS 420
 Db 60 -----SKDRGTGNDEKKA----- 71
 QY 421 DPLIQSSPAVCPPEPPSPKHYSSVTSRTGSSGAKEMKLKGADGKT--KIATPRCAAPP 478
 Db 72 -----KGADGKTGAKIATPRGAAAPA 92
 QY 479 QKQANATRIAPKTPAPKTPPSSGEPKPSGDRSGYSPPGTPGSRSTPSPSLTPPTR 538
 Db 93 QKGTSNATRIAPKTPSPKTPGSGEPKPSGDRSGYSPPGTPGSRSTPSPSLTPPTR 152
 QY 539 EPKKNVAVTPPKSPSSAKSRLOTPAPVMPDLKNVSKIGSTENLKHOPGGKQVQIINKK 598
 Db 153 EPKKNVAVTPPKSPSASKSRLOTPAPVMPDLKNVSKIGSTENLKHOPGGKQVQIINKK 212
 QY 599 LQSLNVQSKGSKDNTHKVPVGGSGVQIYKVPVDLSKVTSCGSLGNIHHKPGGQVEVKS 658
 Db 213 LQSLNVQSKGSKDNTHKVPVGGSGVQIYKVPVDLSKVTSCGSLGNIHHKPGGQVEVKS 272
 QY 659 EKLDKDRVQSKIGSLDNTHVPGGSKNIETHKLTFRENAKAKTDHGAIEIVKSPVVS 718
 Db 273 EKLDKDRVQSKIGSLDNTHVPGGSKNIETHKLTFRENAKAKTDHGAIEIVKSPVVS 332
 QY 719 DTSRHLNVSTGSDIMVDSQPLATLADSVASLAKOGL 758
 Db 333 DTSRHLNVSTGSDIMVDSQPLATLADSVASLAKOGL 372

RESULT 2
 ID Q64715 PRELIMINARY; PRT: 1825 AA.
 AC Q64715;
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
 DE Microtubule-associated protein.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=WISTAR; TISSUE=BRAIN;
 RA Dresse A.E.;
 RL Submitted (JUL-1990) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 1802-1825 FROM N.A.
 RC STRAIN=WISTAR; TISSUE=BRAIN;
 RX MEDLINE=89334524; PubMed=2474284;
 RA Marechal D., Delapierre D., Dresse A.;
 RT "Cloning and partial sequencing of a new rat brain specific cDNA";
 RL Arch. Int. Physiol. Biochim. 96:231-236 (1988).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=WISTAR; TISSUE=BRAIN;
 RX MEDLINE=95094034; PubMed=8000911;
 RA Ferhat L., Ben-Ari Y.P., Khrestchatsky M.;
 RT "Complete sequence of rat MAP2a, a novel MAP2 isoform";
 RL C. R. Acad. Sci., III, Sci. Vie 317:304-309 (1994).
 CC -1- SIMILARITY: CONTAINS 3 TAU/MAP REPEATS.

DR EMBL; X54100; CAA38034.1; -.
 DR EMBL; X74211; CAA52283.1; -.
 DR Pfam; IPR001084; Tubulin_Tau.
 DR Pfam; PF00418; tubulin-binding; 3.
 DR PROSITE; PS00229; TAU_MAP; 2.
 KW Microtubules; Repeat.
 FT CONFLICT 151 1509
 FT CONFLICT 1689 1689
 FT REF. 3).
 SQ SEQUENCE 1825 AA; 198565 MW; 6A920FD9B22213BC CRC64;
 Query Match 20.1%; Score 795; DB 11; Length 1825;
 Best Local Similarity 29.2%; Pred. No. 1.6e-36;
 Matches 275; Conservative 118; Mismatches 273; Indels 276; Gaps 34;
 QY 3 EPROEFVEMDHAGTYGLG-----DRKDOGGYTMHODQEGDT---- 39
 Db 974 EHAKESEVGDKVLFGLGVTYEQTSAKELITYKETAPERAELGLSSVPEVAEVTTKA 1033
 QY 40 DAGL-----KESPLQTPTEGSEEPGSETDAKST-----PTAED- 74
 Db 1034 DOGLDVAAKDDQDPLDIKVSDFQMGASGMSVDAGKTIELKFEVDQQLTSLSEAPQETDS 1093
 QY 75 ---VTAPLVDEGA-----PGQAAAQPHTEIPGTTAEEA--GIGOTPSLEDEA----AG 120
 Db 1094 FMGIESHVKGDAKVSETEVEKVKAPDLVHQEAVDKEESYESSEGESLTMESLKPDGE 1153
 QY 121 HVTQEPESGKVVQGEFLR---EPGPPGLSHQLMS-----GMP-----GAPLL 159
 Db 1154 KKETSPETSLIODEVALKLSVEIPCVPVSEADSSIDEKAEVQMEFIQLPKESTETPDI 1213
 QY 160 PEGPREATR-QPSGTGTPEDTEGGRHAPPELLKQLLDHQBEPPLKAGGKER----- 211
 Db 1214 PAIPSDVTQPOPEAVVSEPAE-----VRGEEIEAEGEVDKLLPRSDT 1257
 QY 212 -----PGSKEEV-----DEDRDVDES---SPQDSPPSKASPAQGRPPQTAAREATS 255
 Db 1258 LQITDLLVPGSREEFVETCPGEHGVVESVVTIEDDFITVVTQTTDEGLGSHSVRFAA- 1316
 QY 256 IPGPPAEGAIPVPDFLSKV--STEIPASEPDGSPVGRAGKQDAPLETFTHVEITPNVOK 313
 Db 1317 -PVQPEERRPYPHDEEVLVMAEAQAEPKDGSDPAPATPEKEEVFFSEYKETYDYDK 1375
 QY 314 EQAHSEELHGRA-----APFGAPGEGPEARGPSLGDETKADLPE 353
 Db 1376 DETTIDDSIMDADSLWVDTDODDRSILTEOLETIPKEERAKEARRPSLEKHKRKE----- 1430
 QY 354 PSEKQPAAPAGKPVSRVPQLKARWVSKSDGTGSD--KAKTSTR 398
 Db 1431 ----KPFKTGRGR--ISTPE---RKVAKKEPSTVSRDEVRKKKAVYKKAELAKKSEVOAH 1481
 QY 399 SSAKTLKNRPCLSPKLPDTPGSSDPLIQSPSPAVCPPEPPSPKHYSSVTSRT----- 449
 Db 1482 SPSRKLILKPAIKYTRPT-----HLSCVKRKTATSGESAQ 1517
 QY 450 GSSGAKEMKLKGADGKTKIATPRGAAP-----PGQKG-----QANATRIAPKTPPAKPT 498
 Db 1518 APSAFKQAKDKVTGKITKSPKRSLLPRSSILPRRGSGDRRENSFSLNSSISSARRT 1577
 QY 499 PPSGEPKPSGDRSGYSPPGSPG---TPGS-----RSRTPSLPTPPT--REP----- 540
 Db 1578 TRS---EPIRRAGKSGTSTPTTPGSTAITPGTPPSPSYSRTPGTGTPSYRTPGTPTKFGIL 1635
 QY 541 ----KQVAVRTPPKSPSSAKSRLOTPAPVMPDLKNVSKIGSTENLKHOPGGKQVQIIN 596
 Db 1636 VPSEKKNVAVTPPKSPATPK-QRLINQPLDPLKNVSKIGSTDNKIKYQPKGQVQIIVT 1694
 QY 597 KKLDSNVQSKGSKDNTHKVPVGGSGVQIYKVPVDLSKVTSCGSLGNIHHKPGGQVEV 656
 Db 1695 KKIDLSH-----VTSKCSGLKNRHRPGGGRVKI 1723
 QY 657 KSEKLDKDRVQSKIGSLDNTHVPGGSKNIETHKLTFRENAKAKTDHGAIEIVKSPV 716
 Db 1723

```
Db 1724 ESVKLDFEKAQAKVGSGLDVAHVPVGGNVKIDSKLNFREHAKARVDHGAIBITQSPSR 1783
QY 717 SGTSPRHLSSVSTGSDVMDVSPQLATLADEVSAKLAKOGL 758
Db 1784 SSVASPRRLSNVSSSGSINLLESPQLATLAEDVTAALAKOGL 1825

RESULT 3
Q63724
ID Q63724 PRELIMINARY; PRT; 1828 AA.
AC Q63724;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Microtubule-associated protein.
OS Rattus rattus (Black rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
ON NCBI_Taxid=10117;
RX SEQUENCE FROM N.A.
RC STRAIN=RA25; TISSUE=BRAIN;
RA Matus A., Doll T.;
RL Submitted (MAY-1990) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 3 TAU/MAP REPEATS.
DR EMBL; X53455; CAA37535.1; -.
DR InterPro; IPR001084; Tubulin_Tau.
DR Pfam; PF00418; tubulin-binding; 3.
DR PROSITE; PS00229; TAU_MAP; 2.
KW Microtubules; Repeat.
SQ SEQUENCE 1828 AA; 198983 MW; 553C3BB5F16D436C CRC64;

Query Match 19.9%; Score 790.5; DB 11; Length 1828;
Best Local Similarity 29.7%; Pred. No. 2.8e-36;
Matches 276; Conservative 121; Mismatches 280; Indels 253; Gaps 35;

QY 3 EPRQEFVEMDHAGTYGLG-----DRKQGGYTMHQDEGDT---- 39
Db 978 EHAKSEEVGDKVFLGLGTYEQTSAKELITTKETAPERAEKGLSSVPEVAEVTYTKA 1037
QY 40 DAGL-----KESPLQTTEDGSEPGSETSDAKST-----PTAED- 74
Db 1038 DQGLDVAARKDQSPDLIDKVSFGQMASGMSVDAGTKIELKFEVDQQLTSLSEAPQETDS 1097
QY 75 ---VTAPLVDEGA----PGKAAQAPHEIPEGITAEAA--GIGDTPSLEDEA---AG 120
Db 1098 FNGIESHVKDKAKVSETVEKVKAPDLVHOEAVDKVESESSGESLTMESLKPDGE 1157
QY 121 HVTQEPESGVVQEGFLR----EPGPPGLSHQMSGMPGAPL-----LPEGPREATROP 170
Db 1158 KKETSPTSLSIQDEVALKLSVEIPCPVPVSEADSSIDEKAEVQMEFIQLPK--EESTETP 1215
QY 171 SGTG-PEDTEGGHAPPELLKHLLGDLHQEGPLPKGAGKER-----PGS 214
Db 1216 DIPAIPSDVT--QPOEALVSE--PAEVRGEEIEIEAGEYDKLLFRSDTLQITDILLVPGS 1272
QY 215 KEEV-----DEDRVDVES---SPQSPSPSKASPAODGRPPQPAAREATSIQCFPAEGAIP 266
Db 1273 REEFVETCGEHKGVVESVVTEDFITVVQTTDDEGLGSHSVRFAA--PVQPEERRP 1330
QY 267 LPVDFLSKV--STEIPASEPDGSPVSGRAKQDAPLEFTFHVITITPNVQKEQAHSEHLGR 324
Db 1331 YPHDELEVLMAEAQAEPKDGSPDAPATPEKEEVPFSEYKTYETDYDYKDETTIDDSIMD 1390
QY 325 A-----AFPGAGEGPEARGPSLGEDTKADLPPEPSKQPAAPR 364
Db 1391 ADSLWVDTODDRSILTEQLETIPKEERAKEARRPSLEKHKR-----KPKFGR 1441
QY 365 GKPVSRVLPOLKARMYSKSDGTGSD-----KKATSTESSAKTLKNRCP 409
Db 1442 GR--ISTPE--RKVAKKEPSTVSRHDEVRKKAIVYKKAELAKKSEVQAHSPSKLILKPA 1496
QY 410 LSPKLPPTGSSDPLIQSPSPAVCPPEPPSPKHKVSSVTSRT-----GSSGAKEMKLIK 460
```

```
Db 1497 IKYTRPT-----HLSVCRKTTATSGESAQAASAKQAKDK 1532
QY 461 GADGKTKTATPRGAAP-----PQKG-----QANATRIAPAKTPAPKPPSSSEPPKSG 509
Db 1533 VTDGITSPEKRSSLPRPSSILPPRRGVSQDRENSFSLSSISSARTTRS--EPIRRA 1590
QY 510 DRSGYSSPGSPG---TPGS---RSRTPSLTPPT--REP-----KKVAVVRT 548
Db 1591 GKSGTSTPTTGTGTAITPGTTPPSYSSRTPTGTPSPVPTPTGTPKSGILVPESEKVAIIRT 1650
QY 549 PPKSPSSAKSLRLQATPVMPDLKNVKSIGSTENLKHQPGGKQVQIINKKLDLNVQSKC 608
Db 1651 PPKSPATPK-QLRLINQPLPDLKNVKSIGSTDNIKYQPKGQVQIVTKKIDLSH----- 1704
QY 609 GSKDNIKHVPGGGVSQIVIKPVDLSKVTSCGSLGNIIHHKPGGQVEVSEKLDKDRVQ 668
Db 1705 -----VTSKCGSLKNIRHRRPGGGRVKIESVKLDFEKAQ 1738
QY 669 SKIGSLDNITHVPGGGNKKIETHKLTFRENAKATDGAETIVKSPVVGDTSPRHLSNV 728
Db 1739 AKVGSLDNAHVPGGGNVKIDSKLNFREHAKARVDHGAETITQSPRSSVSPRLSNV 1798
QY 729 SSTGIDMVDSPQLATLADEVSAKLAKOGL 758
Db 1799 SSSGSINLLESPQLATLAEDVTAALAKOGL 1828

RESULT 4
Q28869
ID Q28869 PRELIMINARY; PRT; 323 AA.
AC Q28869;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Microtubule-associated protein 2 (MAP2) (MAP-2) (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
ON NCBI_Taxid=9913;
RX SEQUENCE FROM N.A., AND CHARACTERIZATION OF TUBULIN REPEATS.
RC TISSUE=BRAIN;
RA MEDLINE=95034751; PubMed=7947727;
RA Coffey R.L., Joly J.C., Cain B.D., Purich D.L.;
RT "Exploring the microtubule-binding region of bovine microtubule-
associated protein-2 (MAP-2): cDNA sequencing, bacterial expression,
and site-directed mutagenesis.";
RL Biochemistry 33:13199-13207(1994).
RN [2]
RP SEQUENCE OF 121-139.
RX MEDLINE=90037224; PubMed=2808529;
RA Joly J.C., Flynn G., Purich D.L.;
RT "The microtubule-binding fragment of microtubule-associated protein-2:
location of the protease-accessible site and identification of an
assembly-promoting peptide.";
RL J. Cell Biol. 109:2289-2294(1989).
RN [3]
RP SEQUENCE OF 18-36 AND 227-258.
RX MEDLINE=92011652; PubMed=1717454;
RA Dingus J., Obar R.A., Hyams J.S., Goedert M., Vallee R.B.;
RT "Use of a heat-stable microtubule-associated protein class-specific
antibody to investigate the mechanism of microtubule binding.";
RL J. Biol. Chem. 266:18854-18860(1991).
RN [4]
RP PHOSPHORYLATION.
RX MEDLINE=98026080; PubMed=9376363;
RA Itoh T.J., Hisanaga S., Hosoi T., Kishimoto T., Hotani H.;
RT "Phosphorylation states of microtubule-associated protein 2 (MAP2)
determine the regulatory role of MAP2 in microtubule dynamics.";
RL Biochemistry 36:12574-12582(1997).
CC -1- FUNCTION: HEAT STABLE PROTEIN WHICH PROMOTES MICROTUBULE ASSEMBLY
```



```

QY 386 TGSDDKKAKTSTRSS-----AKTLKNRPCLSL-----PKLTPGSSDPL 423
Db 446 -----TATQSTSFCEEPDRGIAPEGRPAVSGTNDITPPNKLPLPSPERK 496
QY 424 IQPSSAVCPPEPPSP-KHVSSVTSRTSGGAKEMKLGADGKTKIATPRGAAPQKGO 482
Db 497 TKPIADAKAKRASPSPKASPARSGS-----KSTQTVAKATTASSVASTGPSR 548
QY 483 ANATRIAPAK-----TPAPKTPPSGEPKSGDRSGYSGSPGTPGSRRT 530
Db 549 SPSTLLPKKPTADLSRPKSTSTSMKKTTLTSGTAPAGVAPTRVKP-TP-MPSRPSTP 606
QY 531 SLPTPTTREPKKVAVVTRTPPKSPSSAKSRLQ--TAPVPMPLDLKNVSKIGSTENLKHQP 588
Db 607 SIDKKPT-----SAPSSSTPRLRLATNASAPDLKNVRSKVGSTENLKHQP 654
QY 589 GKGVOIINKLDLSNVQSCGSKDNITKHVPGGSGVOIVYKVDLSKVTSCGSLGNIIHK 648
Db 655 GGRVQIVSKVSYSHIQSCGSKDNITKHVPGGSGVOIQNKVDISKVSKCGSKANIKH 714
QY 649 PGGGQVEVKSEKLDKDRVQSKIGSLDNIHVPGGNGKKTIEHKLTFRENAKAKTDHGAE 708
Db 715 PGGGDVKIESQKLNFEKAQAKVGSJDNVGHLPAGGAVKIETYRLTFRANARATDHGAD 774
QY 709 IYKSPVWSG--DTSPRHLSNVSS 731
Db 775 IVSRPPEPGPGNSGRVLGLSLRT 799

RESULT 6
P97749
ID P97749 PRELIMINARY; PRT; 198 AA.
AC P97749;
DT 01-MAY-1997 (TReMBLrel. 03, Created)
DT 01-MAY-1997 (TReMBLrel. 03, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Microtubule-associated protein (Fragment).
OS Rattus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10118;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=96263369; PubMed=8847742;
RX Vouyiouklis D.A., Brophy P.J.;
RT "Microtubule-associated proteins in developing oligodendrocytes:
RT transient expression of a MAP2c isoform in oligodendrocyte
RT precursors.;"
RL J. Neurosci. Res. 42:803-817(1995).
CC -|- SIMILARITY: CONTAINS 4 TAU/MAP REPEATS.
DR EMBL; S82901; AAB46832.1; -.
DR InterPro; IPR001084; Tubulin_Tau.
DR Pfam; PF00418; tubulin-binding; 4.
DR PROSITE; PS00229; TAU_MAP; 3.
KW Microtubules; Repeat.
FT NON_TER 1
SQ SEQUENCE 198 AA; 21397 MW; 534719226E846078 CRC64;

Query Match
Best Local Similarity 17.4%; Score 690.5; DB 11; Length 198;
Matches 130; Conservative 35; Mismatches 30; Indels 1; Gaps 1;

QY 541 KKVAVVTRPPKSPSSAKSRLQATAPVPMPLDLKNVSKIGSTENLKHQPGGKVOIINKLD 600
Db 4 KKVAIIRTPPKSPATPK-QLRLINQPLDLKNVSKIGSTDNIKYQPGQVRILNKKID 62
QY 601 LSNVQSCGSKDNITKHVPGGSGVOIVYKVDLSKVTSCGSLGNIIHKPGGGQVEVKSEK 660
Db 63 FSKVQSRGSKDNITKHSAGGNGVIVTKKIDLSHVTSCGSLKNIRHRPGGGRVKIESVK 122
QY 661 LDFKDRVQSKIGSLDNIHVPGGNGKKTIEHKLTFRENAKAKTDHGAEIVYKSPVWSGDT 720
Db 720

```

```

Db 123 LDFKEKAQAKVGSJDNVHVPGGNGVKIDSKLNPREHAKARVDHGAEIITQSPSRSSVA 182
QY 721 SPRHLSNVSSSTGSDM 736
Db 183 SPRRLSNVSSSGSINL 198

RESULT 7
Q64710
ID Q64710 PRELIMINARY; PRT; 381 AA.
AC Q64710;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Microtubule-associated protein (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=95046346; PubMed=7957935;
RX Langkopf A., Guilleminot J., Nunez J.;
RT "Two novel HMW MAP2 variants with four microtubule-binding repeats and
RT different projection domains.;"
RL FEBS Lett. 354:259-262(1994).
CC -|- SIMILARITY: CONTAINS 3 TAU/MAP REPEATS.
DR EMBL; S74265; AAB32559.1; -.
DR EMBL; S74266; AAB32560.1; -.
DR InterPro; IPR001084; Tubulin_Tau.
DR Pfam; PF00418; tubulin-binding; 3.
DR PROSITE; PS00229; TAU_MAP; 2.
KW Alternative splicing; Microtubules; Repeat.
FT NON_TER 1
FT VARSPLIC 83 127 MISSING (IN SHORT FORM).
SQ SEQUENCE 381 AA; 41794 MW; ADF65EBF9D35F8D0 CRC64;

Query Match
Best Local Similarity 14.5%; Score 576; DB 11; Length 381;
Matches 146; Conservative 57; Mismatches 97; Indels 78; Gaps 14;

QY 336 EARGPSLGEDTKADLPPESEKQPAAPKPKVSRVPQLKARVMSKDGTSDDKAKT 395
Db 35 EARPSPLEKHKE-----KPFKTCRG----RISTPERKVAKEPSTVSRDEVRKK 81
QY 396 STRSSAKTLKNRPCLSPKLPDTPGSSDPLIQFSPA--VCPEPP---SPKHSVSVTSRT- 449
Db 82 AVYKKAELAK-----ESEVQAHSPSRKILKPAIKYTRPHTLSCVKKRKT 126
QY 450 -----GSSGAKEMKLGADGKTKIATPRGAAP-----PGQKG-----QANATRIAP 490
Db 127 ATGESAQAPSAFKQAKDVTGDKITKSPKRSLPRPSSILPPRRRSGDRENSFSLNS 186
QY 491 KTPPAPKTPPSSGEPKSGDRSGYSGSPGPG---TPGS---RSRTPSLPTPT--REP 540
Db 187 SISSARRTTRS--EPIRRACKSGTSTTPTGTAITGTPPSSSRTPGTPGTPSPYRTP 244
QY 541 -----KKVAVVTRPPKSPSSAKSRLQATAPVPMPLDLKNVSKIGSTENLKHQP 588
Db 245 GTPKSGILVPSSEKVAIIRTPPKSPATPK-QLRLINQPLDLKNVSKIGSTDNIKYQPK 303
QY 589 GKGVOIINKLDLSNVQSCGSKDNITKHVPGGSGVOIVYKVDLSKVTSCGSLGNIIHK 648
Db 304 GQVRIUNKKIDFSKVSQRCGSKDNITKHSAGGNGVIVTKKIDLSHVTSCGSLKNIRH 363
QY 649 PGGGQVEVKSEKLDKDR 666
Db 364 PGGGRVKIESVKLDFKEK 381

RESULT 8
Q9DBX9
ID Q9DBX9 PRELIMINARY; PRT; 414 AA.

```

AC Q9DBX9;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Microtubule-associated protein.
GN MTAP4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=LUNG;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schram L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Moffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Suzuki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.
RA "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
CC -/- SIMILARITY: CONTAINS 4 TAU/MAP REPEATS.
DR EMBL; AK004694; BAB23480.1; -;
DR MGD; MGI:97178; Mtap4.
DR InterPro; IPR002965; P-rich_extensn.
DR IPR001084; Tubulin.Tau.
DR Pfam; PF00418; tubulin-binding; 4.
DR PRINTS; PR01217; PRICHEXTENS.
DR PROSITE; PS00229; TAU_MAP; 3.
KW Microtubules; Repeat.
SQ SEQUENCE 414 AA; 41784 MW; DFB0695FE8FF9D40 CRC64;

Query Match 14.3%; Score 567.5; DB 11; Length 414;
Best Local Similarity 36.7%; Pred. No. 1.7e-24;
Matches 143; Conservative 50; Mismatches 100; Indels 97; Gaps 12;

QY 351 LPEPSEKQAAAPGRKPVSRVP--QLKARMVSKSDGTGSDDKKAKTSTRSSAKTLKLRP 408
DB 8 VPAAPHKRPAAATATARPSTLPARDYKPKPITEAKVA---EKRTSPSKPSSAPALKPGP 63

QY 409 CLSPKLTPTGSSDPLIO--PS--SPA--VCEPPSS-----PKHSSVSTRTGSSAKEM 457
DB 64 KTTPTVSKATSPSTLVSTGSSRSRPAITLTKRPTSIKTEGKPADVRMTAKSASADLSRS 123

QY 458 KLKGADCKTIATPRGAAPQCGQANATRIKATTPATPPSSGPPKSGDRSGYSSP 517
DB 124 KTTSSASVKNRTTPTGAAPP---AGMTSTRV-----KPMASFRSSGALSVDKK----- 169

QY 518 GSPGTGSRSTRPTSLPTTPPREKVAVVRTPPKSPSSAKSRLOQTAPVMPDLKNYKSKI 577
DB 170 -----PT-----STKPSSSAPRVSRLAT--VSAPDLKSVRSKV 201

QY 578 GSTENLKHQPGG-----GKVIINKKL 599
DB 202 GSTENIKHQPGGRKAVEKTEATATTAGKPEPNAVTAAAGSIASAKPPAGKVQIVSKV 261

QY 600 DLSNVQSKGSKDNKIVPGGSGVQIVYKPVLDLSKVTSCGSLGNTLHKPGGQVEVKSE 659
DB 262 SYSHIQKCGSKDNKIVPGGSGVQIVYKPVLDLSKVTSCGSKANIKHKPGGQVIESQ 321

QY 660 KLDFKDRVQSKIGSLDNITHVPGGKNKIE 689
DB 322 KLNFEKAQAQVGLSDNVGHLPGAGGAVKTE 351

RESULT 9
Q8T5L5 PRELIMINARY; PRT; 838 AA.
ID Q8T5L5;
AC Q8T5L5;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Putative microtubule binding protein.
GN 4F11.1.
OS Anopheles gambiae (African malaria mosquito).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
OC Anopheles.
OX NCBI_TaxID=7165;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Thomasova D., Ton L.Q., Collins F.H., Kafatos F.C.;
RT "Sequencing and analysis of Pen1 region from Anopheles gambiae
chromosome 2R."
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ438610; CAD27473.1; -;
SQ SEQUENCE 838 AA; 88130 MW; EAA2EBAEAB04A26 CRC64;

Query Match 12.2%; Score 485; DB 5; Length 838;
Best Local Similarity 24.9%; Pred. No. 1.8e-19;
Matches 208; Conservative 76; Mismatches 269; Indels 282; Gaps 35;

QY 47 PLOTPTEDGSEEGSETSDAKSTPTAEDVTAPLV-----DEG----- 83
DB 106 FSQPTTTRAPRAEVKFSVPLKTPVRLPLPQQOQHQRDTGPALEPAPISHRPP 165

QY 84 -----APGKQAAQHPTEIPEG--TTAEAGIG--DTPSLEDEAAGHVTPQESKV 131
DB 166 PIAHQQAAPFAMDPAHPNPGMPGPMRPPGNGVGPRTGTPT-----QPQPPRPGM 217

QY 132 VOEGFLRPPGPGCLSHQLMGNPGAPLLPEG-----PREATROPSTGTPEDTEGRHAP 185
DB 218 Y-----PQPPGV-----MPMRPQMPGAVPMQPMQPRPPSAOG----- 253

QY 186 ELLKHOLLGLDHLQEGPLKAGGKERP-----GSKEEYDEDRDVEDSSPOSPSKASPA 240
DB 254 -----MQRPMMGQPPPIRPNPMGPPRQI-----SPQNSNLGGMP 292

QY 241 QDCRPPTAAREATSIPIGPAAGAIPLPVFLSKVSTIPIASEPDGPGSVGRAKQDAPLE 300
DB 293 GMVGPPRPPMPQGCAGPGPPQGMRP---NFVNRPMGDPQTSRP--PSGNDNMGGPPPS 347

QY 301 FTFHVEITPNVQKQAEHSEHLGRAAFPCA-----PGEPEAR-----GPSLGEDTK 347
DB 348 -----SATPSVDDDE---DVGIGRLPADNSALNSPNPARAPPNFTMPGPGGIGEREK 399

QY 348 EADLPEPEKOPAAAPRGKPVSRVPOLKARMVSKSDGTGSDDKKAKTSTRSSAKTLKN- 406
DB 400 S-----NPSRPPSVAGSYGKP-----NDHELDSSGRLPLHALKDF 435

QY 407 -----RPLCLSP-KLPTPGSSDPLTOPSSPA----- 430
DB 436 INKEPRPGQSPQSPSGSQSL-----SPANTDENFSYRPGCAKPNSSGQOQOQOQOQY 491

QY 431 -VCEP-----PSSPKHVSVSTRTGSSGAKEMKKGADG-KTKIATPRGAAPGQK- 481
DB 492 KLOPPPGGRPNAPNPSSAVTPTGGR-----AEGDKVTFTQIPNGGGGGGGGGREG 541

QY 482 --QANATRIKATTPAPKTTPSSGPPKSGDR-----SGYSSPGSPGTPGS 525
DB 542 SQEWSRSRPPQOHSMLRTGPKSLAPDHEGDNDSGVDEYTOBKDRPNALASPASP--LKS 599


```
Db 138 -----KPSPP--IPDAPTMEDIAPRELESJNFSETSGTSDQADRI 176
QY 458 KLGADGKTIATPRGAAPPQGOANATRIPAKTPAPKTPPSSG-EPKSGDRSGYSS 516
Db 177 MQNNEERVE-----EKKQMS-----PTPSOPQHKTQORSGIRPPTAILRQPKPI 221
QY 517 PGSPGTPGSRSRTPSLPTPTREPKKVAVVTRTPPKSPSSAKSRLQOTAPVPMDDLK----- 571
Db 222 PA-----SLPRATATPSSQRAISTPRQTAS-----TAPSPRISKMSRER 262
QY 572 -----NVKSGIGSTENLKHQPGGKVOIINKKLDLSNV 604
Db 263 SDVQKSTSTRSIDNVGRMTPKVNAPVNVKSVGVN--HKAGGNVEIFSEK-RLYNA 319
QY 605 QSKCGSKDNKHPVGGSVQIVYKVPVDSL-----KVT----- 636
Db 320 QSKVGLKNATHVAGGNGVQIENRKLDFSAASPQVGSKTNYOPAKSDVKIVSEKLTWQAK 379
QY 637 SKCGLGNIHHKPGGQVEVSEKLDKDRVQSKIGSLDNITHVPGGKNIETHKLTFR 696
Db 380 SKVGSMDNAHAKPAGNVQILSQKLNWK--AESKVGSKDNMNRHPPGGNVQIFDEKIRY- 436
QY 697 ENAKAKTDHGAELVYKSPVVGDTSPRHLS-NVSS 730
Db 437 -----VSTDSSRNHSTLDISS 452
```

Search completed: January 28, 2003, 14:18:26
Job time : 46.3037 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 13:58:08 ; Search time 42.2376 Seconds
(without alignments)
2391.326 Million cell updates/sec

Title: US-09-904-987-5

Perfect score: 3965

Sequence: 1 MAEPQEFVEMDHAGTYGL.....SPQLATLADEVSASLAKOGL 758

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002.*
1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2116.5	53.4	441	15 AAR58810	Human tau protein.
2	2116.5	53.4	441	21 AAY81386	Human paired helic
3	2116.5	53.4	441	23 ABB03398	Human tau protein.
4	2098.5	52.9	441	18 AAE14577	Human Cysless htai
5	2083.5	52.5	441	18 AAW34856	Human tau protein.
6	2083.5	52.5	441	23 ABB05642	Mutant human tau p
7	2077.5	52.4	441	17 AAW05282	Human tau protein.
8	1837.5	46.3	390	17 AAW05283	Truncated human ta
9	1590	40.1	352	19 AAY20248	Human microtubule
10	1586	40.0	351	21 AAY15200	Human Tau protein.

11	1574	39.7	352	10 AAR91294	Paired helical fil
12	1574	39.7	352	14 AAR32708	Human tau-protein.
13	1350.5	34.1	364	23 ABB57300	Mouse ischaemic co
14	1329	33.5	314	19 AAY20956	Human microtubule
15	812.5	20.5	1831	18 AAW23329	Microtubule-associ
16	809.5	20.4	1824	19 AAY20300	Human microtubule
17	807	20.4	1828	23 ABB57249	Mouse ischaemic co
18	801.5	20.2	1717	18 AAW23331	Neuroblastoma indi
19	733	18.5	140	17 AAW05285	Human tau protein
20	657	16.6	125	22 AAB00720	Human htau40 pepti
21	621.5	15.7	1125	21 AAB22934	Mouse microtubule-
22	621.5	15.7	1125	21 AAY79637	Mouse microtubule-
23	621.5	15.7	1610	21 AAB22870	Microtubule associ
24	621.5	15.7	1610	21 AAY79648	EXFP-DEVD-NAP4-EBF
25	593	15.0	112	16 AAR76937	Bifunctional caspa
26	563	14.2	106	17 AAR92516	PHF-tau (143-254)
27	551	13.9	140	17 AAW05286	Microtubule-associ
28	501	12.6	95	17 AAW05284	Fragment of microt
29	471	11.9	700	22 ABB61671	Human tau protein
30	470.5	11.9	94	22 AAB00729	Drosophila melanog
31	462.5	11.7	94	22 AAB00732	Human TAU derived
32	456	11.5	91	22 AAB00731	Human TAU derived
33	455.5	11.5	94	22 AAB00730	Human TAU derived
34	419	10.6	78	22 AAB55640	Tau peptide region
35	364	9.2	67	15 AAR59837	Sequence of human
36	300	7.6	1466	22 AAE02537	Porcine alpha(III)
37	298.5	7.5	1466	22 AAE02533	Bovine alpha(III)
38	298.5	7.5	1466	22 AAE02534	Bovine alpha(III)
39	287.5	7.3	99	21 AAG03156	Human secreted pro
40	285.5	7.2	1806	23 AAG04266	Human endometrial
41	283.5	7.2	1466	22 ABB50291	Collagen type III
42	283.5	7.2	1466	23 ABB50747	Human tumour Endot
43	283.5	7.2	1469	22 ABB51591	Novel human diagno
44	283.5	7.2	1745	22 AAG77793	Human pro-alpha-3(
45	283.5	7.2	1745	23 ABB97234	Novel human protei

ALIGNMENTS

RESULT 1

AAR58810
ID AAR58810 standard; protein; 441 AA.

AC AAR58810;

DT 27-MAR-1995 (first entry)

DE Human tau protein.

DE Tau; cerebrospinal fluid; immunoassay; antibody; detection;

DE diagnosis; central nervous system; CNS; cytopathies; cytopathy;

DE Alzheimer's disease.

OS Homo sapiens.

PN WO9418560-A.

PD 18-AUG-1994.

PF 10-FEB-1994; 94WO-JP00196.

PR 12-FEB-1993; 93JP-0046133.

PR (TEIJ) TEIJIN LTD.

PI Eguchi H, Hosoda K, Kobayashi S, Kubota T, Mori H;

PI Nakamoto T;

PR WPI; 1994-279910/34.

PR Sandwich immunoassay of tau protein in cerebrospinal fluid - for

PT diagnosis of Alzheimer's disease and other CNS cytopathies

```
XX PS Claim 1; Page 16-18; 36pp; Japanese.
XX CC Detection of the human tau protein (or fragments of it) in samples
XX CC of cerebrospinal fluid enables the diagnosis of central nervous
XX CC system cytopathies such as Alzheimer's disease. Detection is
XX CC performed using labelled antibodies which recognise sites within the
XX CC region defined by the amino acid residues 251-441. The antibodies
XX CC are preferably polyclonal.
XX SQ Sequence 441 AA;

Query Match 53.4%; Score 2116.5; DB 15; Length 441;
Best Local Similarity 58.2%; Pred. No. 6.9e-124;
Matches 441; Conservative 0; Mismatches 0; Indels 317; Gaps 2;

Qy 1 MAEPRQEFVEMDHAGTYGLGDRKQGGYTMHQDQGGDTDAGLKESPLQTPPTEDGSEEPG 60
Db 1 MAEPRQEFVEMDHAGTYGLGDRKQGGYTMHQDQGGDTDAGLKESPLQTPPTEDGSEEPG 60
Qy 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAQPHTEIPEGTAAEAGIGDTPSLEDEAAG 120
Db 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAQPHTEIPEGTAAEAGIGDTPSLEDEAAG 120
Qy 121 HVTQEPESGKVVOEGFLREPQPGLSHQLMSGMPGAPLLPEGPREATROPSTGTPEDTEG 180
Db 121 HVTQ----- 124
Qy 181 GRHAPELLKHQLLDLHQEGPLKAGGKERPGSKKEEVEDRDVDDESSPDSPSKASPA 240
Db 125 ----- 124
Qy 241 QDGRPPQTAAREATSIPIGFAEGAIPLVDFLSKVSTEIPASEPDGPGSVGRAKGQDAPLE 300
Db 125 ----- 124
Qy 301 FTFHVEITPNVQKEAHSEHGLRAAFGAPGEGPEARGPSLGEDTKHEADLPPESEKQPA 360
Db 125 ----- 124
Qy 361 AAPRGKPVSRYPQLKARVSKSDGTGDDKAKTSTRSSAKTLKPNRCLSPKLPPTPGSS 420
Db 125 ----- 142
Qy 421 DPLQPSAPVCPBPSPSPKHSVTSRTGSSGAKEMKLGADGKTKIATPRGAAPGQK 480
Db 143 -----KGADGKTKIATPRGAAPGQK 163
Qy 481 GOANATRIPTAPKTPPSSGPPKSGDRSGYSSPGSPGTPGSRSTRPSLPTPTREP 540
Db 164 GOANATRIPTAPKTPPSSGPPKSGDRSGYSSPGSPGTPGSRSTRPSLPTPTREP 223
Qy 541 KKVAVRTPPKSPSAKSRLQTAVPVMDLKNVSKIGSTENLKHQPGGKVQIINKLD 600
Db 224 KKVAVRTPPKSPSAKSRLQTAVPVMDLKNVSKIGSTENLKHQPGGKVQIINKLD 283
Qy 601 LSNVQSCGSKDNKIHVPGGGQVIVYKPVDSLKVTSKCGSLGNHHPGGGQVGVKSEK 660
Db 284 LSNVQSCGSKDNKIHVPGGGQVIVYKPVDSLKVTSKCGSLGNHHPGGGQVGVKSEK 343
Qy 661 LDFKDRVQSKIGSLDNTHVPGGNGKKIETHKLTFRENAKAKTDHGAIEIVKSPVSGDT 720
Db 344 LDFKDRVQSKIGSLDNTHVPGGNGKKIETHKLTFRENAKAKTDHGAIEIVKSPVSGDT 403
Qy 721 SPRHLNSVSTGSDIMVDSPOLATLADSVSASLAKQGL 758
Db 404 SPRHLNSVSTGSDIMVDSPOLATLADSVSASLAKQGL 441

RESULT 2
AAV81386
ID AAV81386 standard; protein; 441 AA.
XX
```

```
AC AAV81386;
XX DT 19-JUN-2000 (first entry)
XX DE Human paired helical filament phosphorylated tau protein.
XX KW Phosphorylated tau protein; human; paired helical filament;
XX KW polyclonal antibody; Alzheimer's disease; cerebrospinal fluid; CSF;
XX KW diagnosis; detection.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT Modified-site 199
FT Modified-site /note= "Phosphorylated"
FT Modified-site 231
FT Modified-site /note= "Phosphorylated"
FT Modified-site 235
FT Modified-site /note= "Phosphorylated"
XX JP20000034300-A.
XX PN 02-FEB-2000.
XX PD
XX PF 17-JUL-1998; 98JP-0204040.
XX PR 17-JUL-1998; 98JP-0204040.
XX PA (MITU ) MITSUBISHI CHEM CORP.
XX DX WPI; 2000-285529/25.
XX PT Anti-phosphated tau protein antibody - for the detection of Alzheimer
XX PS disease
XX CC Claim 1; Page 8-9; 12pp; Japanese.
XX CC The invention relates to an antibody against a phosphorylated tau
XX CC protein (AAV81386), which is a component of the paired helical filament
XX CC found in the plaques associated with Alzheimer's disease. A
XX CC phosphorylated tau protein fragment selected from peptides
XX CC AAV81387-Y81390 is conjugated to keyhole limpet haemocyanin (KLH), and
XX CC used to raise polyclonal antibodies in a rabbit. The antibodies of the
XX CC invention are specific for phosphorylated tau protein and may be used to
XX CC detect phosphorylated tau protein in the cerebrospinal fluid (CSF) of a
XX CC patient suspected of having Alzheimer's disease. Use of the antibodies of
XX CC the invention provides specific diagnosis of Alzheimer's disease. The
XX CC present sequence represents phosphorylated tau protein.
XX SQ Sequence 441 AA;

Query Match 53.4%; Score 2116.5; DB 21; Length 441;
Best Local Similarity 58.2%; Pred. No. 6.9e-124;
Matches 441; Conservative 0; Mismatches 0; Indels 317; Gaps 2;

Qy 1 MAEPRQEFVEMDHAGTYGLGDRKQGGYTMHQDQGGDTDAGLKESPLQTPPTEDGSEEPG 60
Db 1 MAEPRQEFVEMDHAGTYGLGDRKQGGYTMHQDQGGDTDAGLKESPLQTPPTEDGSEEPG 60
Qy 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAQPHTEIPEGTAAEAGIGDTPSLEDEAAG 120
Db 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAQPHTEIPEGTAAEAGIGDTPSLEDEAAG 120
Qy 121 HVTQEPESGKVVOEGFLREPQPGLSHQLMSGMPGAPLLPEGPREATROPSTGTPEDTEG 180
Db 121 HVTQ----- 124
Qy 181 GRHAPELLKHQLLDLHQEGPLKAGGKERPGSKKEEVEDRDVDDESSPDSPSKASPA 240
Db 125 ----- 124
Qy 241 QDGRPPQTAAREATSIPIGFAEGAIPLVDFLSKVSTEIPASEPDGPGSVGRAKGQDAPLE 300
```


Db 125 ----- 124
QY 301 FTFHVEITPNVQKEAHSEHLCRAAPGAPGEGPEARPSLGEDTKADLPPEKQPA 360
Db 125 ----- 124
QY 361 AAPRGKPVSRVQLKARWVSKSDGTGSDDKRAKTSRSSAKTLKRNPCLSPKLTPGSS 420
Db 125 ----- 142
QY 421 DPLIQSSPAVCEPSSPKHVSSTRTGSSGAKEMKLKADGKTKIATPRGAAPGQK 480
Db 143 ----- 163
QY 481 GOANATRIPAKTPPAPKTPSSGEPKSGDRSGYSPGSPGTGSRSTPSPPTREP 540
Db 164 GOANATRIPAKTPPAPKTPSSGEPKSGDRSGYSPGSPGTGSRSTPSPPTREP 223
QY 541 KKVAVRTPKSPSSAKSRLQATAPVMPDLKNVKSIGSTENLKHQPGGKQVQIINKL 600
Db 224 KKVAVRTPKSPSSAKSRLQATAPVMPDLKNVKSIGSTENLKHQPGGKQVQIINKL 283
QY 601 LSNVQSKCGSKONIKHVPGGGVQIYKPVVDSKVTSCGSLGNTHHHPGGQVEVKSEK 660
Db 284 LSNVQSKCGSKONIKHVPGGGVQIYKPVVDSKVTSCGSLGNTHHHPGGQVEVKSEK 343
QY 661 LDFKRVQSKIGSLDNITHVPGGKNIETHKLTFRENAKAKTDHGAIEIVKSPVSGDT 720
Db 344 LDFKRVQSKIGSLDNITHVPGGKNIETHKLTFRENAKAKTDHGAIEIVKSPVSGDT 403
QY 721 SPRHLSNVSTGSDWDSQPLATLADEVASLAKQGL 758
Db 404 SPRHLSNVSTGSDWDSQPLATLADEVASLAKQGL 441

RESULT 3
ID ABB09398
XX ABB09398 standard; peptide: 441 AA.
AC ABB09398;
DT 01-JUL-2002 (first entry)
DE Human Tau protein.
XX
KW NF-H; neurodegenerative disorder; immunogenic epitope; neurotropic;
KW neuroprotective; antiparkinsonian; cerebroprotective; vasotropic;
KW Alzheimer's disease; multi-infarct dementia; Pick's disease;
KW frontotemporal dementia; dementia pugilistica; vascular dementia;
KW Parkinson's disease; Gerstmann-Straussler-Scheinker disease;
KW multiple sclerosis; Amyotrophic lateral sclerosis;
KW transient ischaemic attack; stroke; NF-M; Tau; B-amyloid protein.
OS Homo sapiens.
XX
XX WO200115655-A2.
XX
PD 08-MAR-2001.
XX
PF 27-AUG-2000; 2000WO-IL00509.
XX
PR 31-AUG-1999; 99US-0386347.
PR 27-JUL-2000; 2000US-0221150.
XX
PA (UYRA-) UNIV RAMOT APPLIED RES & IND DEV LTD.
XX
XX Michaelson DW;
PI
XX
DR WPI; 2002-179260/23.
XX
PT Identifying existence, non-existence, type or state of
PT neurodegenerative disorder, by reacting with serum a peptide derived
PT from protein associated with the disorder, to which antibody is

PT produced at onset of disorder -
XX Disclosure; Page 116-117; 117pp; English.
XX
CC The invention relates to identifying existence, non-existence, type or
CC state of neurodegenerative disorder (ND) in individual, involving
CC immunoreacting with serum derived from an individual, a peptide
CC representing an epitope derived from endogenous protein to which an
CC antibody (Ab) is produced in vivo at onset or during progression of ND,
CC where Ab immunobinds with the peptide and detecting presence, absence or
CC degree of immunobinding. Activity of the immunogenic peptides of the
CC invention may be described as, neurotropic, neuroprotective,
CC antiparkinsonian, cerebroprotective and vasotropic. The peptides of the
CC invention are useful for identifying existence, non-existence, type or
CC state of ND in an individual, where ND is associated with progressive
CC loss of cognitive functions or motoric functions, or progressive loss
CC of control of motoric functions. Such diseases include; Alzheimer's
CC disease, Multi-infarct dementia (MID), Pick's disease, frontotemporal
CC dementia, Gerstmann-Straussler-Scheinker disease, Parkinson's
CC sclerosis, Amyotrophic lateral sclerosis (ALS), transient ischaemic
CC attack (TIA) and stroke. In a preferred embodiment of the present
CC invention, the endogenous protein is selected from the group consisting
CC of NF-H, NF-M, tau and B-amyloid protein. The current sequence represents
CC the amino acid sequence of the human Tau protein.
XX
SQ Sequence 441 AA;

Query Match 53.4%; Score 2116.5; DB 23; Length 441;
Best Local Similarity 58.2%; Pred. No. 6.9e-124;
Matches 441; Conservative 0; Mismatches 0; Indels 317; Gaps 2;
QY 1 MAEPQEEFVEMDHAGTYGLGDRKQGGYTMHQDEGDTAGLKESPLQTPTEGSEEPG 60
Db 1 MAEPQEEFVEMDHAGTYGLGDRKQGGYTMHQDEGDTAGLKESPLQTPTEGSEEPG 60
QY 61 SETSDAKSTPTAEDVTAPLVDEGAPGKAAAOPTHTPEGTAAEAGTGDTPSLEDAAG 120
Db 61 SETSDAKSTPTAEDVTAPLVDEGAPGKAAAOPTHTPEGTAAEAGTGDTPSLEDAAG 120
QY 121 HVTQEPESGKVQVQEGFLRPGPGLSHQLMGPGLPPEGPREATRQSGTGPEDEG 180
Db 121 HVTQ----- 124
QY 181 GRHAPELLKHQLLDLHQEGPLKAGGKERPGSKKEEVEDRDVDDESSPQSPSKASPA 240
Db 125 ----- 124
QY 241 QDGRPPQTAAREATSIPOFPAEGAIPLPVDFLSKVSTEIPASEPDGSPVGRAGKQDAPLE 300
Db 125 ----- 124
QY 301 FTFHVEITPNVQKEAHSEHLCRAAPGAPGEGPEARPSLGEDTKADLPPEKQPA 360
Db 125 ----- 124
QY 361 AAPRGKPVSRVQLKARWVSKSDGTGSDDKRAKTSRSSAKTLKRNPCLSPKLTPGSS 420
Db 125 ----- 142
QY 421 DPLIQSSPAVCEPSSPKHVSSTRTGSSGAKEMKLKADGKTKIATPRGAAPGQK 480
Db 143 ----- 163
QY 481 GOANATRIPAKTPPAPKTPSSGEPKSGDRSGYSPGSPGTGSRSTPSPPTREP 540
Db 164 GOANATRIPAKTPPAPKTPSSGEPKSGDRSGYSPGSPGTGSRSTPSPPTREP 223
QY 541 KKVAVRTPKSPSSAKSRLQATAPVMPDLKNVKSIGSTENLKHQPGGKQVQIINKL 600
Db 224 KKVAVRTPKSPSSAKSRLQATAPVMPDLKNVKSIGSTENLKHQPGGKQVQIINKL 283
QY 601 LSNVQSKCGSKONIKHVPGGGVQIYKPVVDSKVTSCGSLGNTHHHPGGQVEVKSEK 660

|||||
Db 284 LSNVQSKGSKDNKIHVPGGSGVQIYKPVDSLKVTSCGSLGNIIHKFGGQGVKSEK 343
QY 661 LDFKDRVQSKIGSLDNIITHVPGGNGKKIETHKLTFFRENAKAKTDHGAETIYKSPVVSQGT 720
Db 344 LDFKDRVQSKIGSLDNIITHVPGGNGKKIETHKLTFFRENAKAKTDHGAETIYKSPVVSQGT 403
QY 721 SPRHLSNVSTGSDIMVDSPLATLADEVSLAKOGL 758
Db 404 SPRHLSNVSTGSDIMVDSPLATLADEVSLAKOGL 441

RESULT 4
AAE14577
ID AAE14577 standard; Protein; 441 AA.
XX
AC AAE14577;
DT 21-MAY-2002 (first entry)
XX
DE Human Cysless htau40 mutant protein.
XX
KW Tau; polymerisation; free fatty acid; FFA; filament; brain;
KW Alzheimer's disease; neurodegenerative disease; Pick's disease;
KW progressive supranuclear palsy; corticobasal degeneration; FTDP-17;
KW frontotemporal dementia and Parkinsonism linked to chromosome 17;
KW site-directed mutagenesis; human; htau40; mutant; muten.
XX
OS Homo sapiens.
OS Synthetic.
PN WO200210770-A2.
XX
PD 07-FEB-2002.
XX
PF 31-JUL-2001; 2001WO-US41512.
XX
PR 31-JUL-2000; 2000US-221777P.
XX
PA (NEUR-) NEURONAUTICS INC.
XX
PI Gamblin TC, Binder LI;
XX
DR WPT; 2002-206223/26.
XX
PT Identifying reagents for ability to effect tau polymerization,
PT comprises altering oxidative conditions for polymerization and/or the
PT oxidative state of a fatty acid used to stimulate polymerization -
XX
PS Example 2; Page 42-43; 56pp; English.
CC The invention relates to a method of identifying reagents that effect
CC tau polymerisation. The method involves adding a free fatty acid (FFA) to
CC a polymerisation buffer to form a FA-containing buffer, adding tau
CC protein to the buffer to form a tau polymerisation reaction mixture,
CC adding the reagent to the reaction mixture, incubating it to form tau
CC filaments, and finally detecting the tau filaments. The invention
CC also provides a method for regulating the assembly of tau protein in the
CC brain of a mammal, by administering an inhibitor of FA oxidation.
CC The method of the invention is useful for screening agents which effect,
CC preferably prevent and/or inhibit, tau polymer formation in the brains of
CC patients with Alzheimer's disease and other related neurodegenerative
CC diseases characterised, at least in part, by tau polymer formation
CC e.g. Pick's disease, progressive supranuclear palsy, corticobasal
CC degeneration and frontotemporal dementia and Parkinsonism
CC linked to chromosome 17 (FTDP-17). The improved protocols of the
CC invention for inducing tau polymerisation utilise modifications of the
CC oxidative environment of the polymerisation reaction or modifications of
CC the oxidative state of components of the in vitro polymerisation
CC reaction. The present sequence is human Cysless htau40 double-mutant
CC protein with c291A and c322A substitutions, used in an
CC exemplification of the invention.
XX

SQ Sequence 441 AA;
Query Match 52.9%; Score 2098.5; DB 23; Length 441;
Best Local Similarity 57.9%; Pred. No. 9.2e-123;
Matches 439; Conservative 0; Mismatches 2; Indels 317; Gaps 2;
QY 1 MAEPQEFVEMDHAGTYGLGDRKDOGGYTHQDQEGDTAGLKESPLQTTEDGSEEPG 60
Db 1 MAEPQEFVEMDHAGTYGLGDRKDOGGYTHQDQEGDTAGLKESPLQTTEDGSEEPG 60
QY 61 SETSDAKSTPTAEDVTAPLVDGAPGKQAAQPHTEIPEGTTAEAGIGDTPSLDEAAG 120
Db 61 SETSDAKSTPTAEDVTAPLVDGAPGKQAAQPHTEIPEGTTAEAGIGDTPSLDEAAG 120
QY 121 HVTQEPESGKVQEGFLREPGLSLHQLMSGMPGAPLLPEGPREATRQPSGTGPEDTEG 180
Db 121 HVTQ----- 124
QY 181 GRHAPELLKHLLGLDHLQEGPLKAGGKERPGSKKEEVEDEDRDVDESSPDSPSKASPA 240
Db 125 ----- 124
QY 241 QDGRPPQTAAREATSIPGPPAEGAIPLPVDLFLSKVSTEIPASEPDGPGSVGRAKQDAPLE 300
Db 125 ----- 124
QY 301 FTFHVEITPVQKEQAHSEHLGRAAFPGAPGEGPEARGPSLGEDTKADLPPESEKQPA 360
Db 125 ----- 124
QY 361 AAPRGKPSRVPLKARMYSKSKDGTGSDDKKAKTSTRSSAKTLKNRCLSPKLTPTGSS 420
Db 125 -----ARMVSKSDGTGSDDKKA----- 142
QY 421 DPLIOPSPAPVCEPPSPKHVSSVTSRTGSGGAKEMKLKGADGKTKIATPRGAAPPQOK 480
Db 143 -----KGADGKTKIATPRGAAPPQOK 163
QY 481 GOANATRIPAKTPAPKTPSSGEPKSGDRSGYSSPGSGPTPGSRSTPSLPTPTREP 540
Db 164 GOANATRIPAKTPAPKTPSSGEPKSGDRSGYSSPGSGPTPGSRSTPSLPTPTREP 223
QY 541 KKVAVVTRTPKSPSSAKSLQATAPVMPDLKNVKSIGSTENLKHQPGGKQVQIINKKLD 600
Db 224 KKVAVVTRTPKSPSSAKSLQATAPVMPDLKNVKSIGSTENLKHQPGGKQVQIINKKLD 283
QY 601 LSNVQSKGSKDNKIKHVPGGSGVQIYKPVDSLKVTSCGSLGNIIHKFGGQGVKSEK 660
Db 284 LSNVQSKGSKDNKIKHVPGGSGVQIYKPVDSLKVTSCGSLGNIIHKFGGQGVKSEK 343
QY 661 LDFKDRVQSKIGSLDNIITHVPGGNGKKIETHKLTFFRENAKAKTDHGAETIYKSPVVSQGT 720
Db 344 LDFKDRVQSKIGSLDNIITHVPGGNGKKIETHKLTFFRENAKAKTDHGAETIYKSPVVSQGT 403
QY 721 SPRHLSNVSTGSDIMVDSPLATLADEVSLAKOGL 758
Db 404 SPRHLSNVSTGSDIMVDSPLATLADEVSLAKOGL 441

RESULT 5
AAW34856
ID AAW34856 standard; protein; 441 AA.
XX
AC AAW34856;
DT 27-MAR-1998 (first entry)
XX
DE Human tau protein.
XX
KW Antibody; phosphorylated tau protein; paired helical filament;
KW detection; Alzheimer's disease; human.
XX
OS Homo sapiens.

```

XX PN W09734145-A1.
XX PD 18-SEP-1997.
XX PF 13-MAR-1997; 97WO-JP00804.
XX PR 13-MAR-1996; 96JP-0056090.
XX PA (MITU) MITSUBISHI CHEM CORP.
XX PI Inahori K, Ishiguro K, Park J, Sato K, Uchida T;
XX DR WPI; 1997-470978/43.
XX PT Antibody prepared using a partial peptide containing part of
XX PT phosphorylated tau protein - used for detecting Alzheimer's disease
XX PS Claim 2; Pages 25-27; 48pp; Japanese.
XX CC An antibody, prepared using a partial peptide containing the
XX CC phosphorylated residue of the phosphorylated tau protein, e.g. the
XX CC present sequence, in a paired helical filament, can be used to
XX CC detect Alzheimer's disease, i.e. by detecting phosphorylated tau
XX CC protein in brain extracts or tissue fragments.
XX SQ Sequence 441 AA;

Query Match 52.5%; Score 2083.5; DB 18; Length 441;
Best Local Similarity 57.8%; Pred. No. 7.9e-122;
Matches 439; Conservative 0; Mismatches 0; Indels 321; Gaps 4;

QY 1 MAEPQEFVEYMEDHAG--TYGLGDRKQGGYTMHQDQEGDTAGLKESPLQPTPTDGSSE 58
DB 1 MAEPQEFVEYMEDHAGQDTYGLGDRKQGGYTMH--QEGDTAGLKESPLQPTPTDGSSE 58

QY 59 PGSETSDAKSTPTAEDVTAPLVDEGAPGKQAAQAQPHTEIPGTTAEAGIGDTPSLEDEA 118
DB 59 PGSETSDAKSTPTAEDVTAPLVDEGAPGKQAAQAQPHTEIPGTTAEAGIGDTPSLEDEA 118

QY 119 AGHVTQEPESGKVVQEGFLREPGLSHQLMSGMPGAPLLPEGPREATRQPSGTGPEDT 178
DB 119 AGHVTQ----- 124

QY 179 EGGRHAPPELLKHOLLGDLHQEGPLKAGGKERPGSKEEVEDRDVDDESSPDSPSKAS 238
DB 125 ----- 124

QY 239 PAODGRPPQTAAREATSIQFPAGAIPLVPDFLSKVSTETPASEPDGSPVGRAGQDAP 298
DB 125 ----- 124

QY 299 LEFTFHEITPNVQKQAHSEHLGRAATPGAPGEPARGPSLGEDTKREADLPEPSEKQ 358
DB 125 ----- 124

QY 359 PAAAPRGKPVSRVPLKARVMVSKDGTGDDKAKTSTRSSAKTLKNRPLCPKLPPTG 418
DB 125 -----ARNVSKDGTGSDDKA----- 142

QY 419 SSOPLQPSPAVCEPPSPKHVSVTSRTSGSSAKEMKLKGADGKTKIATPRGAAPP 478
DB 143 -----KGADGKTKIATPRGAAPP 161

QY 479 QKQANATRIKATPPAPKTPSSGEPKSGDRSGYSSPGSGTGCPSRSTRPSLTPPTR 538
DB 162 QKQANATRIKATPPAPKTPSSGEPKSGDRSGYSSPGSGTGCPSRSTRPSLTPPTR 221

QY 539 EPKKVAVRTPPKSPSSAKSRLOTPAPVMPDLKNVKSKITSTENLKHOPGGKQVLIINK 598
DB 222 EPKKVAVRTPPKSPSSAKSRLOTPAPVMPDLKNVKSKITSTENLKHOPGGKQVLIINK 281

QY 599 LDLSNVQSKCGSKDNKIKHVPGGSVQIYKVPDLSKVTSCGSLGNIHHKPGGQVEVKS 658

```

```

DB 282 LDLSNVQSKCGSKDNKIKHVPGGSVQIYKVPDLSKVTSCGSLGNIHHKPGGQVEVKS 341
QY 659 EKLDKDRVQSKIGSLDNITHVPGGKNKKIETHKLTFRENAKAKTDHGAEIYVKSPVVS 718
DB 342 EKLDKDRVQSKIGSLDNITHVPGGKNKKIETHKLTFRENAKAKTDHGAEIYVKSPVVS 401
QY 719 DTSPRHLSNVSTGSDMDVDSPLATLADSVASLAKOGL 758
DB 402 DTSPRHLSNVSTGSDMDVDSPLATLADSVASLAKOGL 441

RESULT 6
ABB05642
ID ABB05642 standard; Protein; 441 AA.
XX AC ABB05642;
XX DT 25-APR-2002 (first entry)
XX DE Mutant human tau protein SEQ ID NO:9.
XX KW Transgenic mouse; neurofibril; senile dementia; mutant tau protein;
XX KW brain; neurodegenerative disease; Alzheimer's disease; human.
XX OS Homo sapiens.
XX OS Synthetic.
XX PN WO200195709-A1.
XX PD 20-DEC-2001.
XX PF 14-JUN-2001; 2001WO-JP05070.
XX PR 15-JUN-2000; 2000JP-0179723.
XX PA (RIKE) RIKEN KK.
XX PI Takashima A;
XX DR WPI; 2002-097995/13.
XX PT Transgenic mouse with human mutant tau protein, useful in the
XX PT development of treatments for neurodegenerative diseases -
XX PS Claim 2; Page 28-29; 34pp; Japanese.
XX CC The present invention describes a transgenic mouse transformed by human
XX CC mutant tau protein and showing neurofibril changes in the brain due to
XX CC senile dementia. Also described is a method for screening for treatments
XX CC for neurodegenerative diseases. The transgenic mouse can be used in
XX CC developing remedies for neurodegenerative diseases such as Alzheimer's
XX CC disease. The present sequence represents a mutant human tau protein
XX CC which is used in the exemplification of the present invention.
XX SQ Sequence 441 AA;

Query Match 52.5%; Score 2083.5; DB 23; Length 441;
Best Local Similarity 57.8%; Pred. No. 7.9e-122;
Matches 439; Conservative 0; Mismatches 0; Indels 321; Gaps 4;

QY 1 MAEPQEFVEYMEDHAG--TYGLGDRKQGGYTMHQDQEGDTAGLKESPLQPTPTDGSSE 58
DB 1 MAEPQEFVEYMEDHAGQDTYGLGDRKQGGYTMH--QEGDTAGLKESPLQPTPTDGSSE 58

QY 59 PGSETSDAKSTPTAEDVTAPLVDEGAPGKQAAQAQPHTEIPGTTAEAGIGDTPSLEDEA 118
DB 59 PGSETSDAKSTPTAEDVTAPLVDEGAPGKQAAQAQPHTEIPGTTAEAGIGDTPSLEDEA 118

QY 119 AGHVTQEPESGKVVQEGFLREPGLSHQLMSGMPGAPLLPEGPREATRQPSGTGPEDT 178
DB 119 AGHVTQ----- 124

```

QY 179 EGRHAPPELLKHQLLDGDLHQEGPPLKAGAGKRPKSGKEEVEDRDVDESSPODPSPPSKAS 238
 Db 125 ----- 124
 QY 239 PAQDGRPPQTAAREATSIRGFAEGAIPLPVDFLSKVSTEIPASEBDGVSGRAGQDAP 298
 Db 125 ----- 124
 QY 299 LEFTFHEITPNVQKEQAHSEHGLGRAAFPGAGPEGARGPSLGEDTKADLPPEBQ 358
 Db 125 ----- 124
 QY 359 PAAAPRGKPVSRVQLKARWVSKDGTGDDKAKTSTRSSAKTLKNRPCLSPKLPTPG 418
 Db 125 ----- 142
 QY 419 SSDPLIQSPSPVCPPEPPSPKHVSVTSRTGSSGAKEMKLGADGKTKIATPRGAAPP 478
 Db 143 ----- 161
 QY 479 QKQANATRIAPAKTPAPKTPSSGEPKSGDRSGYSSPGSPCTPGSRSTPSLTPPTR 538
 Db 162 QKQANATRIAPAKTPAPKTPSSGEPKSGDRSGYSSPGSPCTPGSRSTPSLTPPTR 221
 QY 539 EPKAVAVRTPPKSPSSAKSRLQTAAPVMPDLKNVSKTGSTENLKHQPGGKQVQIINK 598
 Db 222 EPKAVAVRTPPKSPSSAKSRLQTAAPVMPDLKNVSKTGSTENLKHQPGGKQVQIINK 281
 QY 599 LLSNVQSKCGSKDNTHKHPGGSVQIVYKPVDSLKVTSCKCSLGNHHKPGGGQVEVKS 658
 Db 282 LLSNVQSKCGSKDNTHKHPGGSVQIVYKPVDSLKVTSCKCSLGNHHKPGGGQVEVKS 341
 QY 659 EKLDKDRVQSKIGSLDNTHVPGGKNGKIETHKLTFRENAKAKTDHGAIEIVYKSPVVS 718
 Db 342 EKLDKDRVQSKIGSLDNTHVPGGKNGKIETHKLTFRENAKAKTDHGAIEIVYKSPVVS 401
 QY 719 DTSRHLNSVSTGSDIMVDSPOLATLADEVASLAKQGL 758
 Db 402 DTSRHLNSVSTGSDIMVDSPOLATLADEVASLAKQGL 441

RESULT 7

AAW05282

ID AAW05282 standard; Protein; 441 AA.

XX AAW05282;

XX AC AAW05282;

XX DT 20-DEC-1996 (first entry)

XX DE Human tau protein.

XX KW Tau protein; inhibition; modulation; prophylaxis; treatment;

XX KW Alzheimer's disease; motor neurone disease; Lewy body disease;

XX KW progressive supranuclear palsy; Pick's disease.

XX OS Homo sapiens.

XX PN WO9630766-A1.

XX PD 03-OCT-1996.

XX PF 25-MAR-1996; 96WO-EF01307.

XX PR 27-MAR-1995; 95GB-0006197.

XX PA (HOFF) HOFFMANN LA ROCHE & CO AG F.

XX PI Edwards PC, Harrington CR, Klug A, Roth M, Wischik CM;

XX DR WPI; 1996-455570/45.

XX DR N-PSDB; AAW39591.

XX PT Assay for inhibitors of tau-tau interaction - used for identifying

PT cpds., partic. phenothiazine cpds., for treating pathological
 tau-tau or neuro:filament aggregation

XX Example 2; Page 53-54; 97pp; English.

PS
 XX Detecting an agent which modulates or inhibits tau-tau protein
 association comprises contacting two tau proteins, distinct from
 each other yet capable of binding to the other and where one of the
 tau proteins is labelled, in the presence of the agent suspected of
 being capable of modulating or inhibiting tau-tau interaction.
 Agents identified as being modulators or inhibitors of tau-tau
 interaction may be used for the prophylaxis and treatment of
 Alzheimer's disease, motor neurone disease, Lewy body disease,
 Pick's disease or progressive supranuclear palsy.

SQ Sequence 441 AA;

Query Match 52.4%; Score 2077.5; DB 17; Length 441;

Best Local Similarity 57.5%; Pred. No. 1.9e-121;

Matches 436; Conservative 0; Mismatches 5; Indels 317; Gaps 2;

QY 1 MAEPROBEFVNEDHAGTYGLGDRKDQGYTMHQDQEGDTAGLKESPLQPTEDGSEEP 60

Db 1 MAEPROBEFVNEDHAGTYGLGDRKDQGYTMHQDQEGDTAGLKESPLQPTEDGSEEP 60

QY 61 SETSDAKSTPTAEVDVTAFLVDEGAPGKQAAQAAPHTPEGTAAEEAGIGDTPSLEDAAG 120

Db 61 SETSDAKSTPTAEVDVTAFLVDEGAPGKQAAQAAPHTPEGTAAEEAGIGDTPSLEDAAG 120

QY 121 HVTQEPESGKVQVQEGFLREPGLSHQLMSGMPGAPLLPEGPREATRQPSGTGPEDTEG 180

Db 121 HVTQ----- 124

QY 181 GRHAPPELLKHQLLDGDLHQEGPPLKAGAGKRPKSGKEEVEDRDVDESSPODPSPPSKAS 240

Db 125 ----- 124

QY 241 QDGRPPQTAAREATSIRGFAEGAIPLPVDFLSKVSTEIPASEBDGVSGRAGQDAP 300

Db 125 ----- 124

QY 301 FTFHVEITPNVQKEQAHSEHGLGRAAFPGAGPEGARGPSLGEDTKADLPPESEKQ 360

Db 125 ----- 124

QY 361 AAPRGKPVSRVQPLKARWVSKDGTGDDKAKTSTRSSAKTLKNRPCLSPKLPTPGSS 420

Db 125 ----- 142

QY 421 DPLIQSPSPVCPPEPPSPKHVSVTSRTGSSGAKEMKLGADGKTKIATPRGAAPPQK 480

Db 143 ----- 163

QY 481 QGANATRIAPAKTPAPKTPSSGEPKSGDRSGYSSPGSPCTPGSRSTPSLTPPTR 540

Db 164 QGANATRIAPAKTPAPKTPSSGEPKSGDRSGYSSPGSPCTPGSRSTPSLTPPTR 223

QY 541 KKVAVRTPPKSPSSAKSRLQTAAPVMPDLKNVSKTGSTENLKHQPGGKQVQIINK 600

Db 224 KKVAVRTPPKSPSSAKSRLQTAAPVMPDLKNVSKTGSTENLKHQPGGKQVQIINK 283

QY 601 LSNVQSKCGSKDNTHKHPGGSVQIVYKPVDSLKVTSCKCSLGNHHKPGGGQVEVKS 660

Db 284 LSNVQSKCGSKDNTHKHPGGSVQIVYKPVDSLKVTSCKCSLGNHHKPGGGQVEVKS 343

QY 661 LDFKDRVQSKIGSLDNTHVPGGKNGKIETHKLTFRENAKAKTDHGAIEIVYKSPVVS 720

Db 344 LDFKDRVQSKIGSLDNTHVPGGKNGKIETHKLTFRENAKAKTDHGAIEIVYKSPVVS 403

QY 721 SPRHLNSVSTGSDIMVDSPOLATLADEVASLAKQGL 758

Db 404 SPRHLNSVSTGSDIMVDSPOLATLADEVASLAKQGL 441

RESULT 8

AAW05283
 ID AAW05283 standard; Protein; 390 AA.
 XX
 AC AAW05283;
 XX
 DT 20-DEC-1996 (first entry)
 XX
 DE Truncated human tau protein.
 XX
 KW Tau protein; inhibition; modulation; prophylaxis; treatment;
 KW Alzheimer's disease; motor neurone disease; Lewy body disease;
 KW progressive supranuclear palsy; Pick's disease.
 XX
 OS Homo sapiens.
 XX
 PN W09630766-A1.
 XX
 PD 03-OCT-1996.
 XX
 PF 25-MAR-1996; 96W0-EP01307.
 XX
 PR 27-MAR-1995; 95GB-0006197.
 XX
 PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
 XX
 PI Edwards PC, Harrington CR, Klug A, Roth M, Wischik CM;
 XX
 DR WPI; 1996-455570/45.
 XX
 XX Assay for inhibitors of tau-tau interaction - used for identifying
 PT cpts., partic. phenothiazine cpts.; for treating pathological
 PT tau-tau or neuro:fibrilament aggregation
 XX
 PS Claim 11; : 97pp; English.
 XX
 CC Detecting an agent which modulates or inhibits tau-tau protein
 CC association comprises contacting two tau proteins, distinct from
 CC each other yet capable of binding to the other and where one of the
 CC tau proteins is labelled, in the presence of the agent suspected of
 CC being capable of modulating or inhibiting tau-tau interaction.
 CC Agents identified as being modulators or inhibitors of tau-tau
 CC interaction may be used for the prophylaxis and treatment of
 CC Alzheimer's disease, motor neurone disease, Lewy body disease,
 CC Pick's disease or progressive supranuclear palsy. This sequence of
 CC the human tau protein is truncated at amino acid residue 390. The
 CC full length protein is given in AAW05282.
 XX
 SQ Sequence 390 AA;
 Query Match 46.3%; Score 1837.5; DB 17; Length 390;
 Best Local Similarity 54.6%; Pred. No. 1.4e-106;
 Matches 386; Conservative 0; Mismatches 4; Indels 317; Gaps 2;
 QY 1 MAEPQEFVNEHAGTYGLDRKQGGYTMHODQEGDTAGLKESPLQPTDGESEEPG 60
 DB 1 MAEPQEFVNEHAGTYGLDRKQGGYTMHODQEGDTAGLKESPLQPTDGESEEPG 60
 QY 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAQAQPHTEIPEGTAAEAGIGDTPSLEDAAG 120
 DB 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAQAQPHTEIPEGTAAEAGIGDTPSLEDAAG 120
 QY 121 HVTQEPESGKVVQGFLEPQPGPLSHQLMSGMPGAPLLPEGPREATRQPSGTGPEDETEG 180
 DB 121 HVTQ-----
 QY 181 GRHAPELLKHOLLGDLHQEGPLKAGGKERPGSKKEVEDRDVDESSPDSPPSKASPA 240
 DB 125 -----
 QY 241 QDGRPPOTAAREATSIPTFFAEGAIPLVDFLSKVSTEIPASEPDPGFSVGRAGQDAPLE 300

Db 125 ----- 124
 QY 301 FTFHVEITPNVQEQAHSEHLGRAAFPCAPGEGPEARGPSLGEDTKADLPPESEKQPA 360
 Db 125 ----- 124
 QY 361 AAPRGKPVSRVQLKARMVSKSDGTGSDDKAKKTSTRSSAKTLKNRCLSPKLPPTGSS 420
 Db 125 -----ARWVSKSDGTGSDDKA----- 142
 QY 421 DPLIQSSPAVCPPEPPSSPKHVSVTSRTGSSGAKEMKLGADGKTKIATPRGAAPPQOK 480
 Db 143 -----KGADGKTKIATPRGAAPPQOK 163
 QY 481 GQANATRIPAKTPPAKTPPSSGEPKSGDRSGYSSPGSPGTPGSRSTRPSLPTPTREP 540
 Db 164 GQANATRIPAKTPPAKTPPSSGEPKSGDRSGYSSPGSPGTPGSRSTRPSLPTPTREP 223
 QY 541 KKVAVVTRTPPKSPSSAKSRLQAPVMPDLKNVSKIGSTENLKHQPGGKGVQIINKKLD 600
 Db 224 KKVAVVTRTPPKSLSSAKSRLQAPVMPDLKNVSKIGSTENLKHQPGGKGVQIINKKLD 283
 QY 601 LSNVQSKGSKDNIAKHVPGGGSVQIVYKPVLDLSKVTSCGSLGNTHHKGQGVVEKSEK 660
 Db 284 LSNVQSKGSKDNIAKHVPGGGSVQIVYKPVLDLSKVTSCGSLGNTHHKGQGVVEKSEK 343
 QY 661 LDFKORVQSKIGSLDNITHVPGGSKNIETHKLTIFRENAKAKTDHGA 707
 Db 344 LDFKORVQSKIGSLDNITHVPGGSKNIETHKLTIFRENAKAKTDHGA 390

RESULT 9

AAW05248
 ID AAY20248 standard; Protein; 352 AA.

XX AC AAY20248;

XX DT 22-JUL-1999 (first entry)
 XX DE Human microtubule associated protein Tau wild type fragment.
 XX KW Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
 KW frameshift mutation; age-related disease; neurodegenerative disorder;
 KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
 KW Huntington's disease; multiple sclerosis; alcoholic liver disease;
 KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
 KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
 KW neurofilament-F; presenilin I; presenilin II; cellular tumour antigen;
 KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPP-1;
 KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
 KW high mobility group protein-C; neuroendocrine specific protein A.
 XX OS Homo sapiens.
 XX PN W09845322-A2.
 XX PD 15-OCT-1998.
 XX PF 02-APR-1998; 98W0-IB00705.
 XX PR 10-APR-1997; 97U0-0043163.
 XX PA (UYUT-) RIJKSUNIV UTRECHT.
 XX PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
 XX PA (UYRO-) UNIV ROTTERDAM ERASMUS.
 XX PI Burbach JPH, Grosveld FG, Van Leeuwen FW;
 XX WPI; 1998-609901/51.
 XX DR N-PSDB; AAX75754.
 XX PT Diagnosing disease by detecting frameshift mutations in RNA or
 PT corresponding protein mutations - used to diagnose cancer and

neurological diseases, particularly Alzheimer's disease, and also for treatment and prevention with specific ribozymes or wild-type RNA

Disclosure; Figure 3; 258pp; English.

This invention describes a novel method for the diagnosis of a disease caused by, or associated with, an RNA molecule that has a frameshift mutation. The method is used to diagnose age-related diseases, especially cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's disease, Down's syndrome, myotonic dystrophy, Huntington's disease, multiple sclerosis, alcoholic liver disease, diabetes mellitus type II and many others listed) or susceptibility to these disorders. The method allows a definitive diagnosis of Alzheimer's disease in living patients, at an early stage. It is based on the observation that disease may be caused by mutations in RNA rather than DNA. The invention describes the use of neuronal system RNA molecules, specifically proteins including beta-amyloid precursor protein (beta-APP), the microtubule associated proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule associated protein 2 (MAP2), neurofilament-L, neurofilament-M, neurofilament-F, presenilin I, presenilin II, glial fibrillary acidic protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma 2 (bcl-2) proto-oncogene, semaphorin III, HSPF-1, high mobility group protein-C (HMG-C) and neuroendocrine specific protein A.

Sequence 352 AA;

Query Match 40.1%; Score 1590; DB 19; Length 352;
 Best Local Similarity 46.4%; Pred No. 3.2e-91;
 Matches 352; Conservative 0; Mismatches 0; Indels 406; Gaps 4;

QY 1 MAEPQEFVEMDHACTYGLGRKDGQYTMHQDQEGDTAGLKESPLQTPTDGESEPG 60
 Db 1 MAEPQEFVEMDHACTYGLGRKDGQYTMHQDQEGDTAGLKESPLQTPTDGESEPG 60
 QY 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAQPHTEIPGTTAERAGIGDTPSLDEDAAG 120
 Db 45 -----AEEAGIGDTPSLDEDAAG 62
 QY 121 HVTQEPESGKVVQEGFLREPGLSLHQLSMGMPGAPLLPEGPREATROPSTGTPEDTEG 180
 Db 63 HVTQ----- 66
 QY 181 GRHAPELLKHLLGLDHLQEGPPLKAGGKERPCSKKEVEDRDVDDESSPQSPKASPA 240
 Db 67 ----- 66
 QY 241 QDGRPQTAAREATSTPGFPAECATPLVDFLSKYSTETPASEPDGSPVGRAGQDAPLE 300
 Db 67 ----- 66
 QY 301 FTFPHVEITPNVQEAHSEHLGRAAFPAGGEPGPEARGPSLGEDTKADLPPEPSEKQPA 360
 Db 67 ----- 66
 QY 361 AAPRGKPSRVPOLKARMYVSKSDGTSGDDKAKTSTRSAKTLNRPCLSPKLPPTPGSS 420
 Db 67 -----ARMYVSKSDGTSGDDKKA----- 84
 QY 421 DPLIQSSPAVCPPEPSSPKHYSSVTSRTGSSGAKEMKLKADGKTKATPRGAAPPGOK 480
 Db 85 -----KGADGKTKATPRGAAPPGOK 105
 QY 481 GQANATRIPAKTPAPKTPSSGEPKSGDRSGYSSPGSPGTPGSRSTPSLPTPTPTREP 540
 Db 106 GQANATRIPAKTPAPKTPSSGEPKSGDRSGYSSPGSPGTPGSRSTPSLPTPTPTREP 165
 QY 541 KKVAVRTTPKSPSSAKSLQAPVMPDLKNVSKIGSTENLKHQPGGKGVQIINKLD 600
 Db 165 KKVAVRTTPKSPSSAKSLQAPVMPDLKNVSKIGSTENLKHQPGGK----- 216
 QY 601 LSNVQSCGSKDNKIHVPGGGVSQIVYKPVDSLKVTSCGSLGNHHKPGGQVEVKSEK 660
 Db 601 LSNVQSCGSKDNKIHVPGGGVSQIVYKPVDSLKVTSCGSLGNHHKPGGQVEVKSEK 660

Db 217 -----VOIVYKPVDSLKVTSCGSLGNHHKPGGQVEVKSEK 254
 QY 661 LDFKRVOSKISLDNITHVPGGNGKKIETHKLTFRENAKAKTDHGAEIVYKSPVSGDT 720
 Db 255 LDFKRVOSKISLDNITHVPGGNGKKIETHKLTFRENAKAKTDHGAEIVYKSPVSGDT 314
 QY 721 SPRHLNSVSTGSDIMVDSPOLATLADSVASLAKQGL 758
 Db 315 SPRHLNSVSTGSDIMVDSPOLATLADSVASLAKQGL 352

RESULT 10
 AAY15200
 ID AAY15200 standard; Protein; 351 AA.
 XX
 AC AAY15200;
 XX 28-FEB-2000 (first entry)
 DE Human Tau protein.
 KW Human Tau gene; neurofibrillary tangle formation; abnormal tau filament;
 KW brain; mutation; phosphorylation; isoform ratio; diagnosis; tauopathy;
 KW treatment; neurodegenerative disorder; Fronto-Temporal Dementia;
 KW Familial Multiple System Tauopathy with Presenile Dementia; MSTD;
 KW Pick's Disease; Progressive Supranuclear Palsy; PSP; Alzheimer's Disease;
 KW Corticobasal Degeneration; CB; Prion Protein Cerebral Amyloid Angiopathy;
 KW cognitive disorder.
 XX Homo sapiens.
 OS
 XX
 FH Key Location/Qualifiers
 FT Protein 1..351
 FT /label= Tau_protein
 FT /note= "Microtubule related protein"
 XX
 PN WO9962548-A1.
 XX
 PD 09-DEC-1999.
 XX
 PF 28-MAY-1999; 99WO-US12036.
 XX
 PR 01-JUN-1998; 98US-0087557.
 XX
 PA (ADRE-) ADVANCED RES & TECHNOLOGY INST.
 XX
 PI Ghetti B, Spillantini MG, Murrell JR, Goedert M, Farlow MR;
 PI Klug A;
 XX
 DR WPI; 2000-086858/07.
 XX N-PSDB; AAZ29262.
 PT Diagnosing a tauopathy, especially a Fronto-Temporal Dementia -
 XX
 PS Disclosure; Page 85; 90pp; English.
 XX
 CC The present amino acid sequence is a form of human Tau protein.
 CC There are six tau isoforms, expressed in the normal adult brain with a
 CC slight preponderance of those with 3 repeats over those with 4 repeats.
 CC Mutations in the tau gene affects phosphorylation and leads to formation
 CC of neurofibrillary tangles and alters the tau isoform ratio. The
 CC increased ratio of 4:3 repeat and abnormal tau filaments is closely
 CC related to neurodegenerative disorders. This sequence can be used for
 CC diagnosis and treatment of tauopathies, like Fronto-Temporal Dementia,
 CC Familial Multiple System Tauopathy with Presenile Dementia (MSTD),
 CC Pick's Disease, Progressive Supranuclear Palsy (PSP), Corticobasal
 CC Degeneration (CB) or Alzheimer's Disease. A composition that decreases
 CC the ratio of 4:3 repeat tau isomers, along with an agent for treatment
 CC of a cognitive disorder, is useful for treating a tauopathy. It may also
 CC be useful in diagnosis of Prion Protein Cerebral Amyloid Angiopathy and
 CC other prion protein associated disease characterized by abnormal tau
 CC filament formation.
 XX

XX	PN	W08903993-A.
XX	XX	05-MAY-1989.
XX	XX	19-OCT-1988; 88WO-GB00867.
XX	XX	19-OCT-1987; 87GB-0024412.
XX	XX	(MEDI) MEDICAL RESEARCH COUNCIL.
XX	XX	Wisshik CM, Milstein C, Klug A;
XX	XX	WPI; 1989-150854/20.
XX	XX	Paired helical filament core protein - used for providing reagents sensitive to neurofibrillary tangles used for diagnosing Alzheimer's disease.
XX	XX	Disclosure; fig 1; 29pp; English.
XX	XX	Paired helical filament core protein was sequenced from DNA obtained from brain tissue contg. Alzheimer neurofibrillary tangles. The protein can be used to make Mab's to the PHF core or nucleotide probes, used to diagnose Alzheimer's disease. The protein sequence QIVYKP (Aas 218-223) was used to design the probes.
XX	XX	See also AAN91707.
XX	XX	Sequence 352 AA;
XX	XX	Query Match 39.7%; Score 1574; DB 10; Length 352;
XX	XX	Best Local Similarity 46.2%; Pred. No. 3.2e-90;
XX	XX	Matches 350; Conservative 0; Mismatches 2; Indels 406; Gaps 4;
QY	1	MAEPQEFVEMDHAGTYGLGRKQGGYTHQDQEGTDAGLKESPLQTPTDGSSEPG 60
DB	1	MAEPQEFVEMDHAGTYGLGRKQGGYTHQDQEGTDAGLK- 44
QY	61	SETSDAKSTPTAEDVTAPLVDEGAPGQAAQPHTEIPEGTTAEAGIGDTPSLEDAAG 120
DB	45	-----AEEAGIGDTPSLEDAAG 62
QY	121	HVTQEPESGKVVQEGFLREPGLSHQLMSGMPGAPLLPEGPREATRQSGTGEDTEG 180
DB	63	HVTQ----- 66
QY	181	GRHAPELLKHOLLGDLHQEGPLKAGGKERPGSKKEEVEDRDVDESSPDSPSKASPA 240
DB	67	----- 66
QY	241	QDGRPPQTAAREATSIPIGFPAEGAIPLPVDLFLSKVSTEIPASEPDGVSVGRAKQDAPLE 300
DB	67	----- 66
QY	301	FTFHVETPNVQEAHSEHGLGRAAFPAGPEGPEARGPSLGEDTKADLPPEKQPA 360
DB	67	----- 66
QY	361	AAPRGKPVSRVPOLKARWVSKSDGTGSDDKAKATSTRSSAKTLKRNPCLSPLPTPGSS 420
DB	67	-----ARMVSKSDGTGSDDKA- 84
QY	421	DPLIQSSPAVCPPEPPSPKHSVTSRTGSSGAKEMKLGADGKTKIATPRGAAPGQK 480
DB	85	-----KGADGKTKIATPRGAAPGQK 105
QY	481	GQANATRIPAKTPPAPKTPPSSGEPKSGDRSGYSSPGSPGTGSRSTRPSLTPTPTREP 540
DB	106	GQANATRIPAKTPPAPKTPPSSGEPKSGDRSGYSSPGSPGTGSRSTRPSLTPTPTREP 165
QY	541	KKVAVVTRPPKSPSAKSRLOQTAPVMPDLKNVSKSIGSTENLKHQGGKVVQIINKKLD 600
DB	166	KKVAVVTRPPKSPSAKSRLOQTAPVMPDLKNVSKSIGSTENLKHQGGK----- 216

QY 601 LSNVQSKGSDNIKHVPGGSGYQIVYKPVDSLKSVTSKCGSLGNIHHKPGGGQVEVKSEK 660
 DB 217 -----VQIVYKPVDSLKSVTSKCGSLGNIHHKPGGGQVEVKSEK 254
 QY 661 LDFKDRVQSKIGSLDNITHVPGGNGKKIETHKLTFRENAKAKTDHGAIEIVYKSPVVSQDT 720
 DB 255 LDFKDRVQSKIGSLDNITHVPGGNGKKIETHKLTFRENAKAKTDHGAIEIVYKSPVVSQDT 314
 QY 721 SPRHLSNVSTGSDIMVDSPLATLADEVSASLAKOGL 758
 DB 315 SPRHLSNVSTGSDIMVDSPLATLADEVSASLAKOGL 352

RESULT 12
 AAR32708
 ID AAR32708 standard; Protein; 352 AA.
 XX AC AAR32708;
 XX DT 15-JUN-1993 (first entry)
 XX DE Human tau-protein.
 XX KW Alzheimer's disease; diagnosis; subtyping; monitoring; assay.
 XX OS Homo sapiens.
 XX PN W09303369-A.
 XX PD 18-FEB-1993.
 XX PF 03-AUG-1992; 92WO-0506382.
 XX PR 01-AUG-1991; 91US-073878.
 XX (VOOR/) VOORHEIS P H.
 XX PI Voorheis PH;
 XX DR WPI; 1993-076670/09.
 XX DR N-PSDB; AAQ37305.
 XX PT Method for diagnosing, subtyping and monitoring Alzheimer's
 PT disease - by assaying a sample of body fluid for the presence of a
 PT tau-peptide using an anti-tau antibody
 XX Disclosure; Flg 1; 43pp; English.
 XX The sequence is that one form of human tau protein (from Goedert
 CC et al., PNAS USA 85: 4051-4055) which was used for the prodn.
 CC of anti-tau peptide antibodies. These are used as part of a method
 CC for diagnosing, subtyping or monitoring Alzheimer's disease by
 CC assaying a sample of body fluid for the presence of a tau-peptide
 CC using an anti-tau antibody or the presence of an anti-tau-peptide
 CC autoantibody. The methods can be used for confirming a clinical
 CC diagnosis of Alzheimer's disease and in following the course of the
 CC disease and treatment.
 XX SQ Sequence 352 AA;

Query Match 39.7%; Score 1574; DB 14; Length 352;
 Best Local Similarity 46.2%; Pred. No. 3.2e-90;
 Matches 350; Conservative 0; Mismatches 2; Indels 406; Gaps 4;

QY 1 MAEPQEFVEWEDHAGTYGLGRKDGQGYTMHQDQEGDITDAGLKESPLQTPTEDEGSEPG 60
 DB 1 MAEPQEFVEWEDHAGTYGLGRKDGQGYTMHQDQEGDITDAGLKESPLQTPTEDEGSEPG 44
 QY 61 SETSDAKSTTAEDVTAPLVDEGAPGKQAAQPHTEIPETGTTAEAGIGDTPTPSLEDAAG 120
 DB 45 -----AEEAGIGDTPTPSLEDAAG 62

QY 121 HVTQPESEKVVQEGFLREPGPPGLSHQMGMPAPLLPEGPREATRQPSGTGPEDETEG 180
 DB 63 HVTQ-----
 QY 181 GRHAPELLKHOLLGLHOEGPPLKAGGKERPGSKKEEVEDRDVDESSPQSPKASPA 240
 DB 67 -----
 QY 241 QDGRPPQTAAREATSI PGFPAEGAIPLVDFLSKVSTEIPASEPDGPSVGRAKQDAPLE 300
 DB 67 -----
 QY 301 FTFHVEITPNVQKEAHSEHLLGRAAFPGAPGEGPARGPSLGEDTKREADLPEPSEKQPA 360
 DB 67 -----
 QY 361 AAPRGKPVSRVPLKARMVSKSKDGTGSDDKKAKTSTRSSAKTLKNRPCLSPKLPPTGSS 420
 DB 67 -----ARMVSKSKDGTGSDDKK-----
 QY 421 DFLIQSSPAVCEPPEPPSKHVSSTRTSGSGAKEMKLKCGADGKTKIATPRGAAPGQK 480
 DB 85 -----KCGADGKTKIATPRGAAPGQK 105
 QY 481 GQANATRIPAKTPAPKTPSSGEPKSGDRSGYSSPGSGTSGRSRSTPSLPTPTREP 540
 DB 106 GQANATRIPAKTPAPKTPSSGEPKSGDRSGYSSPGSGTSGRSRSTPSLPTPTREP 165
 QY 541 KKVAVRTPPKSPSAKSRLQTA PVMPDLKNVSKSGISTENLKHQPGGKGVQIINKLKD 600
 DB 166 KKVAVRTPPKSAKSRLQTA PVMPDLKNVSKSGISTENLKHQPGGK----- 216
 QY 601 LSNVQSKGSDNIKHVPGGSGYQIVYKPVDSLKSVTSKCGSLGNIHHKPGGGQVEVKSEK 660
 DB 217 -----VQIVYKPVDSLKSVTSKCGSLGNIHHKPGGGQVEVKSEK 254
 QY 661 LDFKDRVQSKIGSLDNITHVPGGNGKKIETHKLTFRENAKAKTDHGAIEIVYKSPVVSQDT 720
 DB 255 LDFKDRVQSKIGSLDNITHVPGGNGKKIETHKLTFRENAKAKTDHGAIEIVYKSPVVSQDT 314
 QY 721 SPRHLSNVSTGSDIMVDSPLATLADEVSASLAKOGL 758
 DB 315 SPRHLSNVSTGSDIMVDSPLATLADEVSASLAKOGL 352

RESULT 13
 ABB57300
 ID ABB57300 standard; Protein; 364 AA.
 XX AC ABB57300;
 XX DT 07-MAR-2002 (first entry)
 XX DE Mouse ischaemic condition related protein sequence SEQ ID NO:841.
 XX KW Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
 XX vasospastic ischaemia; ischaemic condition; ischaemic disease.
 XX OS Mus musculus.
 XX PN W0200188188-A2.
 XX PD 22-NOV-2001.
 XX PF 18-MAY-2001; 2001WO-JP04192.
 XX PR 18-MAY-2000; 2000JP-0145977.
 XX PA (UYNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.
 XX PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
 XX DR WPI; 2002-034733/04.

CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
 CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
 CC protein-C (HMP-C) and neuroendocrine specific protein A.
 XX
 SQ Sequence 314 AA;

Query Match 33.5%; Score 1329; DB 19; Length 314;
 Best Local Similarity 100.0%; Pred. No. 5.1e-75;
 Matches 251; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 125 EPESKGVQEGFLRPPGGLSHQLMSMPCAPLLPEGPREATRQPSGTGPDTEGGRHA 184

DB 21 EPESKGVQEGFLRPPGGLSHQLMSMPCAPLLPEGPREATRQPSGTGPDTEGGRHA 80

QY 185 PELLKHQLLDLHQBGLKAGGKERPGSKKEEVEDRDVDESSPQSPSKASPAQDGR 244

DB 81 PELLKHQLLDLHQBGLKAGGKERPGSKKEEVEDRDVDESSPQSPSKASPAQDGR 140

QY 245 PPQTAAREATSIIPGPAAGAIPLPVDLFLSKVSTETIPASEPDGSPVGRAGKQDAPLETFH 304

DB 141 PPQTAAREATSIIPGPAAGAIPLPVDLFLSKVSTETIPASEPDGSPVGRAGKQDAPLETFH 200

QY 305 VEITENVQKEQAHSHEHLGRAAFGAPGEGPEARGPSLGEDTKADLPPEPSEKQPAAPR 364

DB 201 VEITENVQKEQAHSHEHLGRAAFGAPGEGPEARGPSLGEDTKADLPPEPSEKQPAAPR 260

QY 365 GKPSRVLPQK 375

DB 261 GKPSRVLPQK 271

RESULT 15

AAW23329

ID AAW23329 standard; protein; 1831 AA.

XX AC AAW23329;

XX DT 23-MAR-1998 (first entry)

XX DE Microtubule-associated protein 2 (MAP2).

XX KW microtubule-associated protein 2; MAP2; neuroblastoma; human;

XX KW isoform; diagnosis; antibody; PCR primer; probe.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Misc-difference 1825

FT Misc-difference 1830

FT Misc-difference 1830

FT Misc-difference 1830

FT Misc-difference 1830

FT Misc-difference 1830

FT Misc-difference 1830

FT Misc-difference 1830

FT Misc-difference 1830

FT Misc-difference 1830

FT Misc-difference 1830

FT Misc-difference 1830

FT Misc-difference 1830

FT Misc-difference 1830

FT Misc-difference 1830

FT Misc-difference 1830

FT Misc-difference 1830

FT Misc-difference 1830

FT Misc-difference 1830

FT Misc-difference 1830

FT Misc-difference 1830

FT Misc-difference 1830

FT Misc-difference 1830

FT Misc-difference 1830

FT Misc-difference 1830

FT Misc-difference 1830

XX This is the human microtubule-associated protein 2 (MAP2). A 250 kD
 CC isoform of this 270 kD MAP2 is indicative of neuroblastoma. This isoform
 CC NB-MAP2 arises from alternative splicing of the primary MAP2 transcripts.
 CC Diagnostic agents for detecting NB-MAP2 in cells can be provided so as
 CC to detect and monitor neuroblastoma. The diagnostic agents are pairs of
 CC primers for specific amplification of DNA corresponding to at least part
 CC of the MAP2 transcript, provided neither primer binds to a sequence
 CC downstream of nucleotide 4168 or upstream of nucleotide 4510 of the
 CC present 5.5 kb sequence, or the corresponding antisense sequence, an
 CC oligonucleotide that binds to NB-MAP2-specific mRNA or the corresponding
 CC DNA in which nucleotides 4168 and 4510 are next to one another, an
 CC antibody specific for an epitope generated by translation of the region
 CC comprising the splice junction of NB-MAP2 specific mRNA, and a cDNA probe
 CC obtained by nick-translation or random priming of DNA or RNA specific for
 CC NB-MAP2 and obtainable from tissue or cell probes. The primers are used to
 CC detect MAP2-specific dimorphisms in tissue and cells by polymerase chain
 CC reaction (PCR) analysis. The oligonucleotide and the probes are used to
 CC detect NB-MAP2 specific transcripts (particularly by binding to membrane-
 CC bound mRNA) and the antibodies are used to detect NB-MAP2 in tissue and
 CC cells. Specifically these tests are used to detect neuroblastoma (or its
 CC precursor cells) and to monitor the success of bone marrow purging. The
 CC probes can detect bone marrow and lung metastases in cases where
 CC histopathological diagnosis of neuroblastoma is difficult. The
 CC oligonucleotide is used therapeutically to inhibit NB-MAP2 expression,
 CC resulting in development of neuroblastoma to normal ganglioneurons.
 XX
 SQ Sequence 1831 AA;

Query Match 20.5%; Score 812.5; DB 18; Length 1831;
 Best Local Similarity 29.2%; Pred. No. 6.8e-42;
 Matches 284; Conservative 120; Mismatches 279; Indels 291; Gaps 39;

QY 5 ROEFVEMDHAGTYGLDRKDDGGYTMHQDQEGDTAGLKESPLQPTPTEDGSEEPGSETS 64

DB 922 KDEFSV-DREASAHISGDK---SGLSKFDEQKKAN-----DRLDVLK-SEEHADSKE 971

QY 65 DAKSTPTAEDVTAPLVDEGAPGKQAAQPHTEIPEGTAAEEA--GIGDTSPSEDEAAGHV 122

DB 972 HAKKTEAGD-EIETFGLVTVTEQALAK-DLSIPTDASSEKAEKGLSSVPEI----- 1021

QY 123 TQPESSKVVQEGF-----LREPGLSLHQLMSGM-----PCAPLLPEGPREA 166

DB 1022 -AEVPSKVEQGLDFAVOGLDVKISDFG-----OMASGLNIDRRRATELKLEATQDM 1074

QY 167 TROPSTGTGPDPT-----EGGRHAPELLKHLQLGD--LHQEGPPLK---GAGGK 209

DB 1075 T--PSSKAPQEAADFMGVESGHMKECTKYSETEVKQKVPDLVHQEAVDKESYESSCE 1132

QY 210 ERPGSKEEVEDRDVDESSPQDS-----PP--SKASPAQDGRP----- 245

DB 1133 HESLTMESLKADEGKETSPESSLIQDEIAVKLSVETPCPPAVSEADLATDERADYQNEF 1192

QY 246 ---PQTAAREATSIIPGPAAGAIPLPVDLFLSKVSTETIPASEPDGSPVGRAGKQDAPLEFT 302

DB 1193 IOGPKEESKETPDISITPDSVAEPLHETIVSE--PAEIQSEEE-----IAQGEYDKLLER 1247

QY 303 F-----HVEITENVQK-----EQAHSEELHGRAAFPG 329

DB 1248 SDTLQITDLGVSGAREFEVETCPSEHKGVIESVVTITDDDFITVWQTTTDEGESGSHVRF 1307

QY 330 APGEGPEARGPSLGEDTKADLPPESEKQ-----PAAAPRGKPSRVVQLKARWVS 380

DB 1308 AALEQPEVERRPSPHDEEFEEFEVEEAQAEPKDGSPASPSPEREVA--LSEYKTKETD 1366

QY 381 KSKDGTGSDS-----KKAKTSS--AKTLKNRP 408

DB 1367 DYKDETTIDDSIMDADSLWVDTDQDDRSIMTEQLETIPKEEAKAEKARRSSLEKHKRKP 1426

QY 409 C-----LSPKPLTPGSSDPL-----IQPSSPA--VCPEPP 436

DB 1427 FKTGRGRISTPERKVAKEPSTVSRDEVRRKAVYKAEKAEKAEKAEKAEKAEKAEKAEK 1486

PT Diagnostic agents for detecting neuroblastoma specific isoform of
 PT microtubule-associated protein - for diagnosis or monitoring of
 PT neuroblastoma, are primers, probes or antibodies, also new
 PT therapeutic antisense sequence

PS Disclosure; Fig 2; 34pp; English.

```

Qy 437 ---SSPKHVSVTSRTSSGA-----KEMKLGADGKTKIATPRGAAP-----PG 478
Db 1487 IKYTRPThLSCVKRKTAAAGGESALAPSVFKQAKDKVSDGVTKSPKRSLSLPRSSILPP 1546
Qy 479 QKG-----QANATRIPAKP-----PAKTPPSS-----GEPPKSG 509
Db 1547 RRGVSGDRDENSFLNSISSARRTTRSEFIRRAGKSGTSTPTPGSTAITFTGTPPSYS 1606
Qy 510 DRGYSSPGSPGTPGSRSTPSLPTPP-----TREPKKAVVTRTPPKSPSSAKSRLQTAP 564
Db 1607 SR---TPGTGCTP-SYPTPTPTGTPKSAIILVPSEKKVAIIRTPKASPGLTPKQLRLIN 1661
Qy 565 VPMDDLNVKSKIGSTENLKHQPGGKVQIINKKLDLSNVQSKGSKDNKIKHVPGGGSVQ 624
Db 1662 QPLPDLKNVSKIGSTONIKYQPKGGQVIVTKKIDLSH----- 1700
Qy 625 IVYKPVDSLKVTSKCSLGNHHPGGGQVEVKSEKLDKDRVQSKIGSLDNITHVPPGG 684
Db 1701 -----VTSKCSLKNIRHRPGGGRVKIESVKLDKFEKAQAKVGSLDNAHHVPPGG 1750
Qy 685 NKIETHKLTRENAKAKTDHGAEIVYKSPVWVGDTSPRHLSNVSSGTSDMVDSPOLAT 744
Db 1751 NVKIDSOKLNFREHAKARVDHGAEIITQSPGRSSVAPRRLSNVSSGSINLLESOLAT 1810
Qy 745 LADEVASLAKQGL 758
Db 1811 LAEDVTAALAKQGL 1824

```

Search completed: January 28, 2003, 14:14:11
Job time : 55.2376 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:19:14 ; Search time 9.93827 Seconds
(without alignments)
1539.035 Million cell updates/sec

Title: US-09-904-987-5

Perfect score: 3965

Sequence: 1 MAEPQREFEVMDHAGTYGL.....SPQLATLADEVSASLAKOGL 758

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 122226 seqs, 20178551 residues

Total number of hits satisfying chosen parameters: 122226

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA.*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	3965	100.0	758	10	US-09-904-987-5
2	2116.5	53.4	441	9	US-10-076-708-2
3	2116.5	53.4	441	9	US-10-066-810-1
4	2077.5	52.4	441	9	US-10-107-181-5
5	1776.5	44.8	383	9	US-10-076-708-4
6	1590	40.1	352	9	US-10-076-708-6
7	1590	40.1	391	10	US-09-790-148-23
8	733	18.5	140	9	US-10-107-181-9
9	575	14.5	109	9	US-10-107-181-3
10	561	14.1	108	9	US-10-107-181-2
11	546.5	13.8	109	9	US-10-107-181-1
12	546	13.8	140	9	US-10-107-181-10
13	364	9.2	67	10	US-09-790-148-1
14	285.5	7.2	1806	10	US-09-919-497-56
15	282	7.1	1325	10	US-09-864-761-35612
16	274	6.9	941	12	US-10-124-557-14
17	274	6.9	1022	12	US-10-124-557-84
18	274	6.9	1038	12	US-10-124-557-74
19	274	6.9	1049	12	US-10-124-557-58

20	274	6.9	1140	12	US-10-124-557-104	Sequence 104, Appl
21	274	6.9	1270	12	US-10-124-557-44	Sequence 44, Appl
22	274	6.9	1311	12	US-10-124-557-42	Sequence 42, Appl
23	274	6.9	1313	12	US-10-124-557-142	Sequence 142, Appl
24	274	6.9	1314	12	US-10-124-557-50	Sequence 50, Appl
25	274	6.9	1320	12	US-10-124-557-46	Sequence 46, Appl
26	274	6.9	1320	12	US-10-124-557-60	Sequence 60, Appl
27	274	6.9	1354	12	US-10-124-557-48	Sequence 48, Appl
28	274	6.9	1361	12	US-10-124-557-40	Sequence 40, Appl
29	274	6.9	1363	12	US-10-124-557-52	Sequence 52, Appl
30	274	6.9	1404	12	US-10-124-557-2	Sequence 2, Appl
31	274	6.9	1404	12	US-10-124-557-62	Sequence 62, Appl
32	247	6.2	1367	10	US-09-801-368-108	Sequence 108, Appl
33	243	6.1	113	10	US-09-864-761-44575	Sequence 44575, A
34	240	6.1	1848	9	US-09-839-996-6	Sequence 6, Appl
35	233	5.9	638	9	US-10-001-887-108	Sequence 108, Appl
36	229	5.8	1907	9	US-09-832-292-39	Sequence 39, Appl
37	228	5.8	2665	10	US-09-864-761-34248	Sequence 34248, A
38	227	5.7	826	10	US-09-894-998-47	Sequence 47, Appl
39	225.5	5.7	617	10	US-09-864-761-36182	Sequence 36182, A
40	223.5	5.6	1162	10	US-09-745-008-34	Sequence 34, Appl
41	222.5	5.6	881	10	US-09-816-860A-2	Sequence 2, Appl
42	221	5.6	755	10	US-09-919-497-57	Sequence 57, Appl
43	218	5.5	503	9	US-10-078-547-2	Sequence 2, Appl
44	218	5.5	802	10	US-09-823-240-2	Sequence 2, Appl
45	216.5	5.5	507	9	US-10-078-547-24	Sequence 24, Appl

ALIGNMENTS

RESULT 1

US-09-904-987-5
; Sequence 5, Application US/09904987
; Patent No. US20020037908A1
; GENERAL INFORMATION:
; APPLICANT: No. US20020037908A1actyl, Inc.
; TITLE OF INVENTION: Methods and Compositions for Controlling Pathological and Prep
; TITLE OF INVENTION: Protein Assembly or Aggregation
; FILE REFERENCE: 42108/26146
; CURRENT APPLICATION NUMBER: US/09/904,987
; CURRENT FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 758
; TYPE: PRT
; ORGANISM: homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NCBI ENTREZ / NM_016835
; DATABASE ENTRY DATE: 2001-02-13
; RELEVANT RESIDUES: (1)..(758)
; US-09-904-987-5

Query Match	100.0%	Score 3965;	DB 10;	Length 758;
Best Local Similarity	100.0%	Pred. No. 5.6e-180;		
Matches 758;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY 1	MAEPQREFEVMDHAGTYGLGDRKDQGGYTMHQDEGDTAGLKESPLQPTTDEGSEPG 60			
Db				
1	MAEPQREFEVMDHAGTYGLGDRKDQGGYTMHQDEGDTAGLKESPLQPTTDEGSEPG 60			
QY 61	SETSDAKSTPTAEDYATPLVDEGAPGKQAAAQPHTEIPEGTTAEAGTGDTPSLEDEAAG 120			
Db				
61	SETSDAKSTPTAEDYATPLVDEGAPGKQAAAQPHTEIPEGTTAEAGTGDTPSLEDEAAG 120			
QY 121	HYVQPESGKVVOEGFLREPGLSHOLMSGMPCAPLLPEGPREATRQPSGTGPDTEG 180			
Db				
121	HYVQPESGKVVOEGFLREPGLSHOLMSGMPCAPLLPEGPREATRQPSGTGPDTEG 180			
QY 181	GRHAPPELLKHQLLDLHOGDPLKAGGKERPGSKKEEVEDRDVDESSPQSPFSKASPA 240			
Db				
181	GRHAPPELLKHQLLDLHOGDPLKAGGKERPGSKKEEVEDRDVDESSPQSPFSKASPA 240			

```
QY 241 ODGRPPQTAAREATSIPIGFAEAGIPLPVDFLSKVSTEIPASEPDGSPVGRAGQDAPLE 300
DB 241 ODGRPPQTAAREATSIPIGFAEAGIPLPVDFLSKVSTEIPASEPDGSPVGRAGQDAPLE 300
QY 301 FTFHVEITPNVQEQAHSEHGLRAAFPGAPGEPARGPSLGEDTKADLPPESEKQPA 360
DB 301 FTFHVEITPNVQEQAHSEHGLRAAFPGAPGEPARGPSLGEDTKADLPPESEKQPA 360
QY 361 AAPRGKPVSRVLPOLKARMVSKSDGTGSDDKAKAKTSTRSSAKTLKNRCLSPKLPPTGSS 420
DB 361 AAPRGKPVSRVLPOLKARMVSKSDGTGSDDKAKAKTSTRSSAKTLKNRCLSPKLPPTGSS 420
QY 421 DPLIOPSSPAVCPEPPSPKHVSSVTSRTGSSGAKEMKLGADGKTKIATPRGAAPPGOK 480
DB 421 DPLIOPSSPAVCPEPPSPKHVSSVTSRTGSSGAKEMKLGADGKTKIATPRGAAPPGOK 480
QY 481 GOANATRIPAKTPAPKTPSSGEPKSGDRSGYSGPGSGTSGRSRTPSLPTPTREP 540
DB 481 GOANATRIPAKTPAPKTPSSGEPKSGDRSGYSGPGSGTSGRSRTPSLPTPTREP 540
QY 541 KKVAVVRTPPKSPSSAKSRLQOTAPVMPDLKNVSKIGSTENLKHQPGGKQVQIINKKLD 600
DB 541 KKVAVVRTPPKSPSSAKSRLQOTAPVMPDLKNVSKIGSTENLKHQPGGKQVQIINKKLD 600
QY 601 LSNVQSKGSKDNKIHVPVGGGQVQIYKPVLDLSKVTSCGSLGNIHHKPGGQVEVYKSEK 660
DB 601 LSNVQSKGSKDNKIHVPVGGGQVQIYKPVLDLSKVTSCGSLGNIHHKPGGQVEVYKSEK 660
QY 661 LDFKDRVQSKIGSLDNIHVPVGGGKNIETHKLTFRENAKAKTDHGAEIVYKSPVWSGDT 720
DB 661 LDFKDRVQSKIGSLDNIHVPVGGGKNIETHKLTFRENAKAKTDHGAEIVYKSPVWSGDT 720
QY 721 SPRHLSNVSTGSDIMVDSPLATLADEVSAKQGL 758
DB 721 SPRHLSNVSTGSDIMVDSPLATLADEVSAKQGL 758

RESULT 2
US-10-076-708-2
; Sequence 2, Application US/10076708
; Patent No. US20020164657A1
; GENERAL INFORMATION:
; APPLICANT: Sharma, Satish
; APPLICANT: Rank, Kenneth
; TITLE OF INVENTION: Assays for Accessing Aa-Tau Aggregation
; FILE REFERENCE: 6322
; CURRENT APPLICATION NUMBER: US/10/076,708
; CURRENT FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 2
; LENGTH: 441
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-076-708-2

Query Match 53.4%; Score 2116.5; DB 9; Length 441;
Best Local Similarity 58.2%; Pred. No. 3.1e-93;
Matches 441; Conservative 0; Mismatches 0; Indels 317; Gaps 2;

QY 1 MAEPQREFEYVMEHDAGTYGLGDRKDGQGYTMHQDQEGDTDAGLKESPLQPTPTDGESEEPG 60
DB 1 MAEPQREFEYVMEHDAGTYGLGDRKDGQGYTMHQDQEGDTDAGLKESPLQPTPTDGESEEPG 60
QY 61 SETSDAKSTPTAEDVTAPLVDEGAPKQAAQPHTEIPGTTAEAEAGIDGTPSLEDEAAG 120
DB 61 SETSDAKSTPTAEDVTAPLVDEGAPKQAAQPHTEIPGTTAEAEAGIDGTPSLEDEAAG 120
QY 121 HVTQEPESGKVVOEGFLREPGLPSHLQLSMGMPGAPLLPEGPREATRQPSGTGPDTEG 180
DB 121 HVTQ----- 124
```

```
QY 181 GRHAPELLKHQLGLHOGEPPLKGAGGKERPGSKKEEYDEDRVDDESSPDSPSKASPA 240
DB 125 ----- 124
QY 241 ODGRPPQTAAREATSIPIGFAEAGIPLPVDFLSKVSTEIPASEPDGSPVGRAGQDAPLE 300
DB 125 ----- 124
QY 301 FTFHVEITPNVQEQAHSEHGLRAAFPGAPGEPARGPSLGEDTKADLPPESEKQPA 360
DB 125 ----- 124
QY 361 AAPRGKPVSRVLPOLKARMVSKSDGTGSDDKAKAKTSTRSSAKTLKNRCLSPKLPPTGSS 420
DB 125 ----- 142
QY 421 DPLIOPSSPAVCPEPPSPKHVSSVTSRTGSSGAKEMKLGADGKTKIATPRGAAPPGOK 480
DB 143 -----KGADGKTKIATPRGAAPPGOK 163
QY 481 GOANATRIPAKTPAPKTPSSGEPKSGDRSGYSGPGSGTSGRSRTPSLPTPTREP 540
DB 164 GOANATRIPAKTPAPKTPSSGEPKSGDRSGYSGPGSGTSGRSRTPSLPTPTREP 223
QY 541 KKVAVVRTPPKSPSSAKSRLQOTAPVMPDLKNVSKIGSTENLKHQPGGKQVQIINKKLD 600
DB 224 KKVAVVRTPPKSPSSAKSRLQOTAPVMPDLKNVSKIGSTENLKHQPGGKQVQIINKKLD 283
QY 601 LSNVQSKGSKDNKIHVPVGGGQVQIYKPVLDLSKVTSCGSLGNIHHKPGGQVEVYKSEK 660
DB 284 LSNVQSKGSKDNKIHVPVGGGQVQIYKPVLDLSKVTSCGSLGNIHHKPGGQVEVYKSEK 343
QY 661 LDFKDRVQSKIGSLDNIHVPVGGGKNIETHKLTFRENAKAKTDHGAEIVYKSPVWSGDT 720
DB 344 LDFKDRVQSKIGSLDNIHVPVGGGKNIETHKLTFRENAKAKTDHGAEIVYKSPVWSGDT 403
QY 721 SPRHLSNVSTGSDIMVDSPLATLADEVSAKQGL 758
DB 404 SPRHLSNVSTGSDIMVDSPLATLADEVSAKQGL 441

RESULT 3
US-10-066-810-1
; Sequence 1, Application US/10066810
; Publication No. US20020188106A1
; GENERAL INFORMATION:
; APPLICANT: Mandelkow, Eva-Maria
; Mandelkow, Eckhard
; Lichtenberg-Kraag, Birgit
; Biernat, Jacek
; Drewes, Gerard
; Steiner, Barbara
; TITLE OF INVENTION: No. US20020188106A1el Tools For The Diagnosis And Treatment Of Alzheimer's Disease
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Tape
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/066,810
; FILING DATE: 04-Feb-2002
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Joseph A. Williams, Jr.
```

REGISTRATION NUMBER: 38,659
 REFERENCE/DOCKET NUMBER: 28384/32777
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312-474-6300
 TELEFAX: 312-484-0448
 TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 441 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 US-10-066-810-1

Query Match 53.4%; Score 2116.5; DB 9; Length 441;
 Best Local Similarity 58.2%; Pred. No. 3.1e-93;
 Matches 441; Conservative 0; Mismatches 0; Indels 317; Gaps 2;

QY 1 MAEPROFEFVEMDHAGTYGLGDRKDGQGYTMHQDQEGDTPDAGLKESPLQTPPTEDGSEEPG 60
 Db 1 MAEPROFEFVEMDHAGTYGLGDRKDGQGYTMHQDQEGDTPDAGLKESPLQTPPTEDGSEEPG 60

QY 61 SETSDAKSTPTAEDVTAPLVDEGAPKQAAQPHTEIPGTTAEAGIGDTPSLDEAAG 120
 Db 61 SETSDAKSTPTAEDVTAPLVDEGAPKQAAQPHTEIPGTTAEAGIGDTPSLDEAAG 120

QY 121 HVTQEPESKVVQEGFLREPGLSHQLMSGMPGAPLLPEGPREATROPSTGTPEDTEG 180
 Db 121 HVTQ----- 124

QY 181 GRHAPELLKHQLLDLHQEGPPLKGAGGKERPGSKKEEVEDRDVDSESPQSPSKASPA 240
 Db 125 ----- 124

QY 241 QDGRPPQTAAREATSIPGPAEGAIPLPVDFLSKVSTEIPASEPDGPGSVGRAGKQDAPLE 300
 Db 125 ----- 124

QY 301 FTFHVEITPNVQKEQAHEHLGRAAFPAGCEGPEARGPSLGEDTKKADLPPEPSEKQPA 360
 Db 125 ----- 124

QY 361 AAPRGKPSRVLPQLKARVSKSDGTGDDKAKTSTRSSAKTLKNRCLSPKLPPTPGSS 420
 Db 125 -----ARVSKSDGTGDDKKA----- 142

QY 421 DPLIQSPSPAVCPPEPPSPKHVSVTSRTGSSGAKEMKLKGADGKTKIATPRGAAPPQOK 480
 Db 143 -----KGADGKTKIATPRGAAPPQOK 163

QY 481 GOANATRIAPAKTPPAPKTPSSGEPKSGDRSGYSSPGTTPGSRSTPSLPTPTREP 540
 Db 164 GOANATRIAPAKTPPAPKTPSSGEPKSGDRSGYSSPGTTPGSRSTPSLPTPTREP 223

QY 541 KKVAVRTTPPKSPSSAKSRLOTAPVMPDLKNVSKIGSTENLKHQPGSGKVQIINKKLD 600
 Db 224 KKVAVRTTPPKSPSSAKSRLOTAPVMPDLKNVSKIGSTENLKHQPGSGKVQIINKKLD 283

QY 601 LSNVSKGSKDNKIKHVPGGGSQIVYKPVDSLKVTSCGSLGNIIHHPGGGQVEVSKSEK 660
 Db 284 LSNVSKGSKDNKIKHVPGGGSQIVYKPVDSLKVTSCGSLGNIIHHPGGGQVEVSKSEK 343

QY 661 LDFKDRVQSKIGSLNIIHVPGGGNKKIETHKLTFRENAKAKTDHGAIVYKSPVWSDT 720
 Db 344 LDFKDRVQSKIGSLNIIHVPGGGNKKIETHKLTFRENAKAKTDHGAIVYKSPVWSDT 403

QY 721 SPRHLSNVSSSTGSDIMWDSPOLATLADEVASLAKOGL 758
 Db 404 SPRHLSNVSSSTGSDIMWDSPOLATLADEVASLAKOGL 441

RESULT 4

US-10-107-181-5
 ; Sequence 5, Application US/10107181
 ; Patent No. US20020168687A1
 ; GENERAL INFORMATION:
 ; APPLICANT: HOFFMANN-LA ROCHE AG
 ; TITLE OF INVENTION: INHIBITION OF TAU-TAU-ASSOCIATION
 ; NUMBER OF SEQUENCES: 12
 ; CORRESPONDENCE ADDRESS:
 ; STREET: Grenzachstrasse 124
 ; CITY: Basle
 ; STATE: BS
 ; COUNTRY: Switzerland
 ; ZIP: CH-4070
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/10107.181
 ; FILING DATE: 28-Mar-2002
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/913.915
 ; FILING DATE: <Unknown>
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 061-688 51 08
 ; TELEFAX: 061-688 13 95
 ; TELEX: 962292/965542 hlr ch
 ; INFORMATION FOR SEQ ID NO: 5:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 441 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
 US-10-107-181-5

Query Match 52.4%; Score 2077.5; DB 9; Length 441;
 Best Local Similarity 57.5%; Pred. No. 2.1e-91;
 Matches 436; Conservative 0; Mismatches 5; Indels 317; Gaps 2;

QY 1 MAEPROFEFVEMDHAGTYGLGDRKDGQGYTMHQDQEGDTPDAGLKESPLQTPPTEDGSEEPG 60
 Db 1 MAEPROFEFVEMDHAGTYGLGDRKDGQGYTMHQDQEGDTPDAGLKESPLQTPPTEDGSEEPG 60

QY 61 SETSDAKSTPTAEDVTAPLVDEGAPKQAAQPHTEIPGTTAEAGIGDTPSLDEAAG 120
 Db 61 SETSDAKSTPTAEDVTAPLVDEGAPKQAAQPHTEIPGTTAEAGIGDTPSLDEAAG 120

QY 121 HVTQEPESKVVQEGFLREPGLSHQLMSGMPGAPLLPEGPREATROPSTGTPEDTEG 180
 Db 121 HVTQ----- 124

QY 181 GRHAPELLKHQLLDLHQEGPPLKGAGGKERPGSKKEEVEDRDVDSESPQSPSKASPA 240
 Db 125 ----- 124

QY 241 QDGRPPQTAAREATSIPGPAEGAIPLPVDFLSKVSTEIPASEPDGPGSVGRAGKQDAPLE 300
 Db 125 ----- 124

QY 301 FTFHVEITPNVQKEQAHEHLGRAAFPAGCEGPEARGPSLGEDTKKADLPPEPSEKQPA 360
 Db 125 -----ARVSKSDGTGDDKKA----- 142

QY 421 DPLIQSPSPAVCPPEPPSPKHVSVTSRTGSSGAKEMKLKGADGKTKIATPRGAAPPQOK 480
 Db 143 -----KGADGKTKIATPRGAAPPQOK 163

QY 481 GOANATRIPAKTPAPKTPSSGEPKSGDRSGYSSPGSGTSGSRSTPSLPTPTREP 540
Db 164 QGANATRIIPAKTPAPKTPSSGEPKSGDRSGYSSPGSGTSGSRSTPSLPTPTREP 223
QY 541 KKVAVRTTPKSPSSAKSRLQATAPVMPDLKNVSKIGSTENLKHQPGGKQVQIINKKLD 600
Db 224 KKVAVRTTPKSLSSAKSRLQATAPVMPDLKNVSKIGSTENLKHQPGGKQVQIINKKLD 283
QY 601 LSNVQSKGSKONIKHVPGGGSGVQIYKPVDSLKVTSKCGSLGNTHHKPGGQVEVKSEK 660
Db 284 LSNVQSKGSKONIKHVPGGGSGVQIYKPVDSLKVTSKCGSLGNTHHKPGGQVEVKSEK 343
QY 661 LDFKDRVQSKIGSLDNITHVPGGNGKKIETHKLTFRENAKAKTDHGAELVYKSPVVS GDT 720
Db 344 LDFKDRVQSKIGSLDNITHVPGGNGKKIETHKLTFRENAKAKTDHGAELVYKSPVVS GDT 403
QY 721 SPRHLSNVSTSGSIDMVDSPQLATLADEVSASLAKOGL 758
Db 404 SPRHLSNVSTSGSIDMVDSPQLATLADEVSASLAKOGL 441

RESULT 5

US-10-076-708-4

; Sequence 4, Application US/10076708
; Patent No. US20020164657A1
; GENERAL INFORMATION:
; APPLICANT: Sharma, Satish
; APPLICANT: Rank, Kenneth
; TITLE OF INVENTION: Assays for Accessing Aa-Tau Aggregation
; FILE REFERENCE: 6322
; CURRENT APPLICATION NUMBER: US/10/076,708
; CURRENT FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 4
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-076-708-4

Query Match 44.8%; Score 1776.5; DB 9; Length 383;
Best Local Similarity 50.5%; Pred. No. 2.7e-77;
Matches 383; Conservative 0; Mismatches 0; Indels 375; Gaps 3;

QY 1 MAEPQREFEVMEDHAGTYGLGDRKDQGGYTMHQDQEGDTDAGLKESPLQTPPTDGSSEPG 60
Db 1 MAEPQREFEVMEDHAGTYGLGDRKDQGGYTMHQDQEGDTDAGLKESPLQTPPTDGSSEPG 44
QY 61 SETSDAKSTPTAEDVTAPLVDEGAPKQAAQPHTEIPEGTTAEAGIGDTPSLEDEAAG 120
Db 45 SETSDAKSTPTAEDVTAPLVDEGAPKQAAQPHTEIPEGTTAEAGIGDTPSLEDEAAG 62
QY 121 HVTQEPESGKVQEGFLREPGLSHQLSHMGMPGAPLLPEGPREATROPSTGTPEDTEG 180
Db 63 HVTQEPESGKVQEGFLREPGLSHQLSHMGMPGAPLLPEGPREATROPSTGTPEDTEG 66
QY 181 GRHAPELLKHQLLDLHOGEPPLKAGGKERPGSKEEVEDRDVDESSPDSPPSKASPA 240
Db 67 GRHAPELLKHQLLDLHOGEPPLKAGGKERPGSKEEVEDRDVDESSPDSPPSKASPA 66
QY 241 QDGRPPQTAAREATSIPGPAEGAIPLPVDFLSKVSTEIPASEPDGPGSVGRAKQDAPLE 300
Db 67 QDGRPPQTAAREATSIPGPAEGAIPLPVDFLSKVSTEIPASEPDGPGSVGRAKQDAPLE 66
QY 301 FTFHVEITPNVQEQAHSEHLGRAAFPAPGAPGEPARGSLGDTKEADLPPESEKQPA 360
Db 67 FTFHVEITPNVQEQAHSEHLGRAAFPAPGAPGEPARGSLGDTKEADLPPESEKQPA 66
QY 361 AAPRGKPVSRVQPKARVSKSKDGTGSDDKKAKTSTRSSAKTLKNRCLSPKLPPTGSS 420
Db 67 AAPRGKPVSRVQPKARVSKSKDGTGSDDKKAKTSTRSSAKTLKNRCLSPKLPPTGSS 84

QY 421 DPLIQSSPAVCEPPSPKHSVSTRTGSSGAKEMKLGADGKTKIATPRGAAPPQK 480
Db 85 DPLIQSSPAVCEPPSPKHSVSTRTGSSGAKEMKLGADGKTKIATPRGAAPPQK 105
QY 481 GOANATRIIPAKTPAPKTPSSGEPKSGDRSGYSSPGSGTSGSRSTPSLPTPTREP 540
Db 106 GOANATRIIPAKTPAPKTPSSGEPKSGDRSGYSSPGSGTSGSRSTPSLPTPTREP 165
QY 541 KKVAVRTTPKSPSSAKSRLQATAPVMPDLKNVSKIGSTENLKHQPGGKQVQIINKKLD 600
Db 166 KKVAVRTTPKSPSSAKSRLQATAPVMPDLKNVSKIGSTENLKHQPGGKQVQIINKKLD 225
QY 601 LSNVQSKGSKDNIKHVPGGGSGVQIYKPVDSLKVTSKCGSLGNTHHKPGGQVEVKSEK 660
Db 226 LSNVQSKGSKDNIKHVPGGGSGVQIYKPVDSLKVTSKCGSLGNTHHKPGGQVEVKSEK 285
QY 661 LDFKDRVQSKIGSLDNITHVPGGNGKKIETHKLTFRENAKAKTDHGAELVYKSPVVS GDT 720
Db 286 LDFKDRVQSKIGSLDNITHVPGGNGKKIETHKLTFRENAKAKTDHGAELVYKSPVVS GDT 345
QY 721 SPRHLSNVSTSGSIDMVDSPQLATLADEVSASLAKOGL 758
Db 346 SPRHLSNVSTSGSIDMVDSPQLATLADEVSASLAKOGL 383

RESULT 6

US-10-076-708-6

; Sequence 6, Application US/10076708
; Patent No. US20020164657A1
; GENERAL INFORMATION:
; APPLICANT: Rank, Kenneth
; APPLICANT: Sharma, Satish
; TITLE OF INVENTION: Assays for Accessing Aa-Tau Aggregation
; FILE REFERENCE: 6322
; CURRENT APPLICATION NUMBER: US/10/076,708
; CURRENT FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 6
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-076-708-6

Query Match 40.1%; Score 1590; DB 9; Length 352;
Best Local Similarity 46.4%; Pred. No. 1.5e-68;
Matches 352; Conservative 0; Mismatches 0; Indels 406; Gaps 4;

QY 1 MAEPQREFEVMEDHAGTYGLGDRKDQGGYTMHQDQEGDTDAGLKESPLQTPPTDGSSEPG 60
Db 1 MAEPQREFEVMEDHAGTYGLGDRKDQGGYTMHQDQEGDTDAGLKESPLQTPPTDGSSEPG 44
QY 61 SETSDAKSTPTAEDVTAPLVDEGAPKQAAQPHTEIPEGTTAEAGIGDTPSLEDEAAG 120
Db 45 SETSDAKSTPTAEDVTAPLVDEGAPKQAAQPHTEIPEGTTAEAGIGDTPSLEDEAAG 62
QY 121 HVTQEPESGKVQEGFLREPGLSHQLSHMGMPGAPLLPEGPREATROPSTGTPEDTEG 180
Db 63 HVTQEPESGKVQEGFLREPGLSHQLSHMGMPGAPLLPEGPREATROPSTGTPEDTEG 66
QY 181 GRHAPELLKHQLLDLHOGEPPLKAGGKERPGSKEEVEDRDVDESSPDSPPSKASPA 240
Db 67 GRHAPELLKHQLLDLHOGEPPLKAGGKERPGSKEEVEDRDVDESSPDSPPSKASPA 66
QY 241 QDGRPPQTAAREATSIPGPAEGAIPLPVDFLSKVSTEIPASEPDGPGSVGRAKQDAPLE 300
Db 67 QDGRPPQTAAREATSIPGPAEGAIPLPVDFLSKVSTEIPASEPDGPGSVGRAKQDAPLE 66
QY 301 FTFHVEITPNVQEQAHSEHLGRAAFPAPGAPGEPARGSLGDTKEADLPPESEKQPA 360
Db 67 FTFHVEITPNVQEQAHSEHLGRAAFPAPGAPGEPARGSLGDTKEADLPPESEKQPA 66
QY 361 AAPRGKPVSRVQPKARVSKSKDGTGSDDKKAKTSTRSSAKTLKNRCLSPKLPPTGSS 420

Db 67 -----ARMVSKSDGTGSDDKA----- 84
QY 421 DPLIQSSPAVCEPPSSPKHVSSTVTSRGSSGAKEMKLGADGKTKIATPRGAAPPGOK 480
Db 85 -----KGADGKTKIATPRGAAPPGOK 105
QY 481 GOANATRIAPKTPPSSGEPKSGDRSGYSSPGSGTTPGSRSTPSLPTPTREP 540
Db 106 GOANATRIAPKTPPSSGEPKSGDRSGYSSPGSGTTPGSRSTPSLPTPTREP 165
QY 541 KKVAVRTPPKSSAKSRLQATAPVMPDLKNVSKIGSTENLKHQPGGKQVQIINKKLD 600
Db 166 KKVAVRTPPKSSAKSRLQATAPVMPDLKNVSKIGSTENLKHQPGGKQVQIINKKLD 216
QY 601 LSNVQSKGSDNKHVPGGSGVQIYVYKPVDSLKVTSCGSLGNTHHKGPGGQVEVYKSEK 660
Db 217 -----VQIYVYKPVDSLKVTSCGSLGNTHHKGPGGQVEVYKSEK 254
QY 661 LDFKDRVQSKIGSLDNITHVPGGSKKIEIETHKLTFFRENAKAKTDHGAEIYKSPVVSQGT 720
Db 255 LDFKDRVQSKIGSLDNITHVPGGSKKIEIETHKLTFFRENAKAKTDHGAEIYKSPVVSQGT 314
QY 721 SPRHLNSVSTGSDIMVDSPOLATLADEVASLAKOGL 758
Db 315 SPRHLNSVSTGSDIMVDSPOLATLADEVASLAKOGL 352

RESULT 7

US-09-790-148-23
; Sequence 23, Application US/09790148
; Patent No. US20020001857A1
; GENERAL INFORMATION:

APPLICANT: VANDERMEEREN, MARC;
MERCKEN, MARC;
VANMECHELEN, EUGEN;
VAN DE VOORDE, ANDRE

TITLE OF INVENTION: MONOCLONAL ANTIBODIES

DIRECTED AGAINST THE MICROTUBULE-ASSOCIATED
PROTEIN TAU, HYBRIDOMAS SECRETING THESE
ANTIBODIES, ANTIGEN RECOGNITION BY THESE

NUMBER OF SEQUENCES: 24

CORRESPONDENCE ADDRESS:

ADDRESSEE: BIERMAN & MUSERLIAN
STREET: 600 THIRD AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10016

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/790,148
FILING DATE: 21-Feb-2001

PRIOR APPLICATION DATA:

APPLICATION NUMBER: <Unknown>
FILING DATE: 27-JUN-1994
APPLICATION NUMBER: 08/256,167
FILING DATE: 13-JUN-1994
APPLICATION NUMBER: 08/244,951
FILING DATE: 10-DEC-1993
APPLICATION NUMBER: PCT/EP93/03499
FILING DATE: 14-DEC-1992
APPLICATION NUMBER: EP/92/403403.6
FILING DATE: 14-DEC-1992

ATTORNEY/AGENT INFORMATION:

NAME: CHARLES A. MUSERLIAN
REGISTRATION NUMBER: 19,683
REFERENCE/DOCKET NUMBER: 410,003-1
TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 661-8000
; TELEFAX: (212) 661-8002
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 391
; TYPE: Amino Acid
; STRANDEDNESS: Unknown
; TOPOLOGY: Unknown
; SEQUENCE DESCRIPTION: SEQ ID NO: 23:
US-09-790-148-23

Query Match 40.1%; Score 1590; DB 10; Length 391;
Best Local Similarity 46.4%; Pred. No. 1.6e-68;
Matches 352; Conservative 0; Mismatches 0; Indels 406; Gaps 4;

QY 1 MAEPQEFVEMDHAGTYGLGDRKDQGGYTMHQDQGGTDAGLKESPLQPTEDGSEEPG 60
Db 40 MAEPQEFVEMDHAGTYGLGDRKDQGGYTMHQDQGGTDAGLK----- 83
QY 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAAQPHTEIPEGTTAEAGIGDTPSLEDEAAG 120
Db 84 -----AEEAGIGDTPSLEDEAAG 101
QY 121 HVTQEPESGKVVQEGFLREPGLSHQMSGMPGAPLLPEGPREATRQPSGTGPDTEG 180
Db 102 HVTQ----- 105
QY 181 GRHAPELLKHOLLGDLHQEGPPLKGAGGKERPGSKEEVEDRDVEDSPDPPSKASPA 240
Db 106 ----- 105
QY 241 QDGRPPQTAAREATSI PGPAEGAIPLPVDFLSKVSTEIPASEPDGVSVGRAKGQDAPLE 300
Db 106 ----- 105
QY 301 FTFHVEITPNVQEQAHSEHGRAAPGAPGEGPEARGPSLGEDTKEADLPESKQPA 360
Db 106 ----- 105
QY 361 AAPRGKPVSRVQDLKARMVSKSDGTGSDDKKAKTSTRSSAKTLKNRPLCLSPKLPFGSS 420
Db 106 -----ARMVSKSDGTGSDDKKA----- 123
QY 421 DPLIQSSPAVCEPPSSPKHVSSTVTSRGSSGAKEMKLGADGKTKIATPRGAAPPGOK 480
Db 124 -----KGADGKTKIATPRGAAPPGOK 144
QY 481 GOANATRIAPKTPPSSGEPKSGDRSGYSSPGSGTTPGSRSTPSLPTPTREP 540
Db 145 GOANATRIAPKTPPSSGEPKSGDRSGYSSPGSGTTPGSRSTPSLPTPTREP 204
QY 541 KKVAVRTPPKSSAKSRLQATAPVMPDLKNVSKIGSTENLKHQPGGKQVQIINKKLD 600
Db 205 KKVAVRTPPKSSAKSRLQATAPVMPDLKNVSKIGSTENLKHQPGGK----- 255
QY 601 LSNVQSKGSDNKHVPGGSGVQIYVYKPVDSLKVTSCGSLGNTHHKGPGGQVEVYKSEK 660
Db 256 -----VQIYVYKPVDSLKVTSCGSLGNTHHKGPGGQVEVYKSEK 293
QY 661 LDFKDRVQSKIGSLDNITHVPGGSKKIEIETHKLTFFRENAKAKTDHGAEIYKSPVVSQGT 720
Db 294 LDFKDRVQSKIGSLDNITHVPGGSKKIEIETHKLTFFRENAKAKTDHGAEIYKSPVVSQGT 353
QY 721 SPRHLNSVSTGSDIMVDSPOLATLADEVASLAKOGL 758
Db 354 SPRHLNSVSTGSDIMVDSPOLATLADEVASLAKOGL 891

RESULT 8

US-10-107-181-9
; Sequence 9, Application US/10107181
; Patent No. us20020168687A1
; GENERAL INFORMATION:

```

, REFERENCE NUMBER: 28-Mar-2002
, FILING DATE: 28-Mar-2002
, PRIORITY:
, PRIOR APPLICATION DATA:
, APPLICATION NUMBER: US/08/913,915
, FILING DATE: <Unknown>
, TELECOMMUNICATION INFORMATION:
, TELEPHONE: 061-688 51 08
, TELEFAX: 061-688 13 95
, TELEX: 962292/965542 hlr ch
, INFORMATION FOR SEQ ID NO: 2:
, SEQUENCE CHARACTERISTICS:
, LENGTH: 108 amino acids
, TYPE: amino acid
, STRANDEDNESS: <Unknown>
, TOPOLOGY: linear
, MOLECULE TYPE: peptide
, SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-107-181-2

Query Match
Best Local Similarity 14.1%; Score 561; DB 9; Length 108;
Matches 108; Conservative 100.0%; Pred. No. 1.le-20;
Mismatches 0; Indels 0;

```

QY 569 DLKNNVSKIGSTENLKHQPGGKVQIINKKLDLSNVQSKGSKDNKIKHVPGGSVQIYVK 628
 Db 1 DLKNNVSKIGSTENLKHQPGGKVQIINKKLDLSNVQSKGSKDNKIKHVPGGSVQIYVK 60
 QY 629 PVDLSKVTSCGSLGNIHHKPGGGQVEVSEKLDKDRVQSKIGSLDN 676
 Db 61 PVDLSKVTSCGSLGNIHHKPGGGQVEVSEKLDKDRVQSKIGSLDN 108

RESULT 11

US-10-107-181-1
 ; Sequence 1, Application US/10107181
 ; Patent No. US20020168687A1
 ; GENERAL INFORMATION:
 ; APPLICANT: HOFFMANN-LA ROCHE AG
 ; TITLE OF INVENTION: INHIBITION OF TAU-TAU-ASSOCIATION
 ; NUMBER OF SEQUENCES: 12
 ; CORRESPONDENCE ADDRESS:
 ; STREET: Grenzacherstrasse 124
 ; CITY: Basle
 ; STATE: BS
 ; COUNTRY: Switzerland
 ; ZIP: CH-4070

COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/10/107,181
 ; FILING DATE: 28-Mar-2002
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/913,915
 ; FILING DATE: <Unknown>
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 061-688 51 08
 ; TELEFAX: 061-688 13 95
 ; TELEX: 962292/965542 hlr ch
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 109 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: <Unknown>
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-10-107-181-1
 Query Match 13.8%; Score 546.5; DB 9; Length 109;
 Best Local Similarity 77.9%; Pred. No. 5.5e-20;
 Matches 109; Conservative 0; Mismatches 0; Indels 31; Gaps 1;
 QY 569 DLKNNVSKIGSTENLKHQPGGKVQIINKKLDLSNVQSKGSKDNKIKHVPGGSVQIYVK 628
 Db 1 DLKNNVSKIGSTENLKHQPGGKVQIINKKLDLSNVQSKGSKDNKIKHVPGGSVQIYVK 29
 QY 629 PVDLSKVTSCGSLGNIHHKPGGGQVEVSEKLDKDRVQSKIGSLDNITHVPGGGNKKI 688
 Db 30 PVDLSKVTSCGSLGNIHHKPGGGQVEVSEKLDKDRVQSKIGSLDNITHVPGGGNKKI 89
 QY 689 ETHKLTFRENAKAKTDHGAE 708
 Db 90 ETHKLTFRENAKAKTDHGAE 109

RESULT 12

US-10-107-181-10
 ; Sequence 10, Application US/10107181
 ; Patent No. US20020168687A1
 ; GENERAL INFORMATION:
 ; APPLICANT: HOFFMANN-LA ROCHE AG
 ; TITLE OF INVENTION: INHIBITION OF TAU-TAU-ASSOCIATION
 ; NUMBER OF SEQUENCES: 12

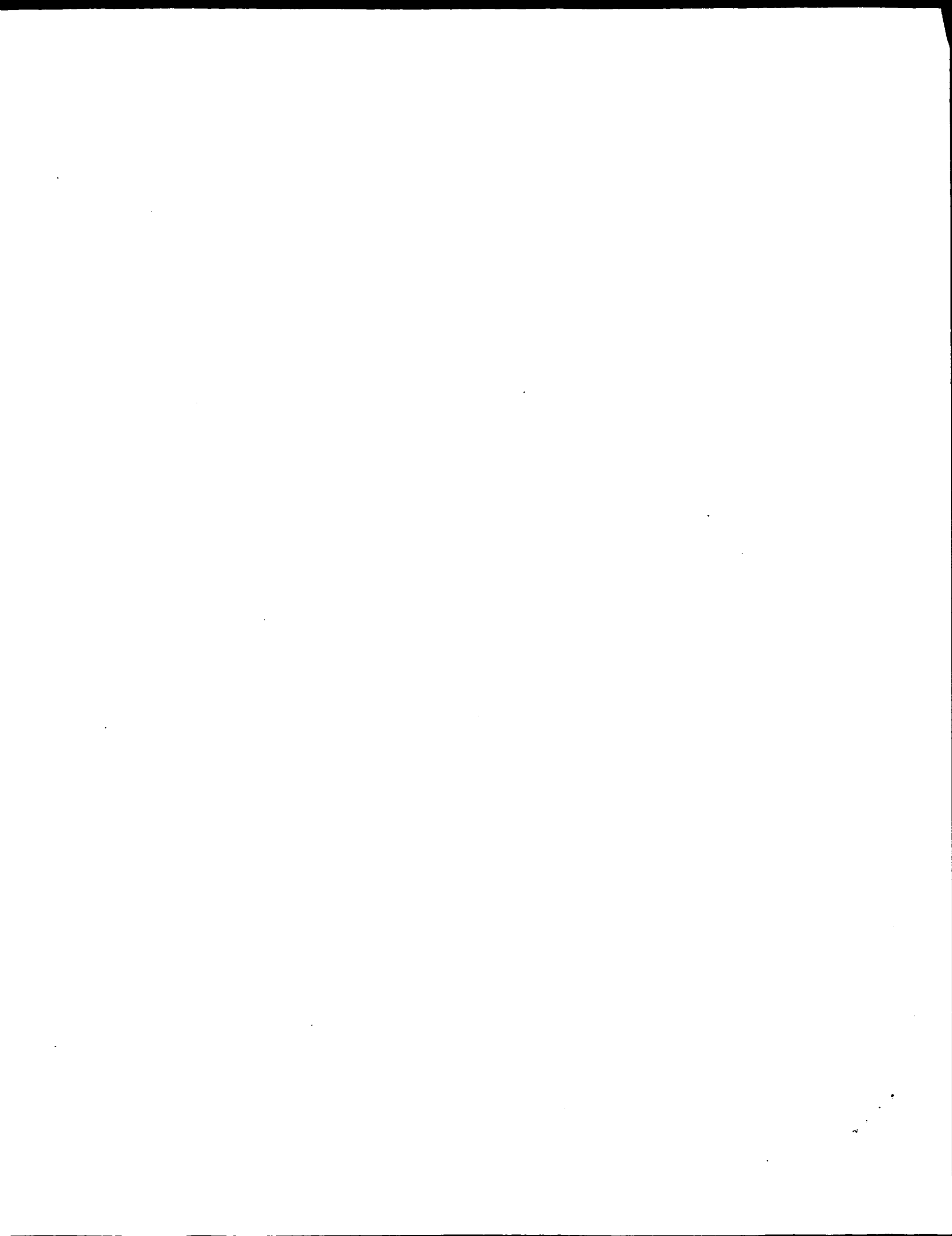
CORRESPONDENCE ADDRESS:
 ; STREET: Grenzacherstrasse 124
 ; CITY: Basle
 ; STATE: BS
 ; COUNTRY: Switzerland
 ; ZIP: CH-4070
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/10/107,181
 ; FILING DATE: 28-Mar-2002

PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/913,915
 ; FILING DATE: <Unknown>
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 061-688 51 08
 ; TELEFAX: 061-688 13 95
 ; TELEX: 962292/965542 hlr ch
 ; INFORMATION FOR SEQ ID NO: 10:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 140 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: <Unknown>
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 10:
 ; US-10-107-181-10

Query Match 13.8%; Score 546; DB 9; Length 140;
 Best Local Similarity 71.4%; Pred. No. 7.6e-20;
 Matches 100; Conservative 23; Mismatches 17; Indels 0; Gaps 0;
 QY 569 DLKNNVSKIGSTENLKHQPGGKVQIINKKLDLSNVQSKGSKDNKIKHVPGGSVQIYVK 628
 Db 1 DLKNNVSKIGSTENLKHQPGGKVQIINKKLDLSNVQSKGSKDNKIKHVPGGSVQIYVK 60
 QY 629 PVDLSKVTSCGSLGNIHHKPGGGQVEVSEKLDKDRVQSKIGSLDNITHVPGGGNKKI 688
 Db 61 KIDLSHVTSCGSLGNIHHKPGGGQVEVSEKLDKDRVQSKIGSLDNITHVPGGGNKKI 120
 QY 689 ETHKLTFRENAKAKTDHGAE 708
 Db 121 DSQKLNFRHAKARVDHGAE 140

RESULT 13

US-09-790-148-1
 ; Sequence 1, Application US/09790148
 ; Patent No. US20020001857A1
 ; GENERAL INFORMATION:
 ; APPLICANT: VANDERMEEREN, MARC;
 ; MERCKEN, MARC;
 ; VANHEHELEN, EUGEN;
 ; VAN DE VOORDE, ANDRE
 ; TITLE OF INVENTION: MONOCLONAL ANTIBODIES
 ; DIRECTED AGAINST THE MICROTUBULE-ASSOCIATED
 ; PROTEIN TAU, HYBRIDOMAS SECRETING THESE
 ; ANTIBODIES, ANTIGEN RECOGNITION BY THESE
 ; NUMBER OF SEQUENCES: 24
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: BIERMAN & MUSERLIAN
 ; STREET: 600 THIRD AVENUE
 ; CITY: NEW YORK
 ; STATE: NEW YORK
 ; COUNTRY: USA
 ; ZIP: 10016
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: FLOPPY DISK
 ; COMPUTER: IBM PC COMPATIBLE
 ; OPERATING SYSTEM: PC-DOS/MS-DOS



GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:14:29 ; Search time 196.488 Seconds
(without alignments)
2487.216 Million cell updates/sec

Title: US-09-904-987-5
Perfect score: 3965
Sequence: 1 MAEPQFEVMEHAGTYGL.....SPQLATLADEVSASLAKGL 758

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues
Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_Main : *
1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US097_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US099_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3965	100.0	758	1 PCT-US02-22032-5	Sequence 5, Appli
2	3965	100.0	758	23 US-09-904-987-5	Sequence 5, Appli
3	3965	100.0	758	27 US-60-406-385-228	Sequence 228, App
4	2116.5	53.4	441	8 US-08-477-648-1	Sequence 1, Appli
5	2116.5	53.4	441	8 US-08-484-674-1	Sequence 1, Appli
6	2116.5	53.4	441	14 US-09-035-708-1	Sequence 1, Appli

7	2116.5	53.4	441	14	US-09-035-708A-1	Sequence 1, Appli
8	2116.5	53.4	441	17	US-09-386-347-71	Sequence 71, Appli
9	2116.5	53.4	441	20	US-09-640-737-1	Sequence 1, Appli
10	2116.5	53.4	441	20	US-09-640-738-1	Sequence 1, Appli
11	2116.5	53.4	441	24	US-10-066-810-1	Sequence 2, Appli
12	2116.5	53.4	441	24	US-10-076-708-2	Sequence 2, Appli
13	2116.5	53.4	441	27	US-60-271-102-2	Sequence 1, Appli
14	2083.5	52.5	441	15	US-09-142-613-1	Sequence 1, Appli
15	2077.5	52.4	441	25	US-10-107-181-5	Sequence 5, Appli
16	1776.5	44.8	383	24	US-10-076-708-4	Sequence 4, Appli
17	1776.5	44.8	383	27	US-60-271-102-4	Sequence 4, Appli
18	1590	40.1	352	24	US-10-076-708-6	Sequence 6, Appli
19	1590	40.1	352	27	US-60-271-102-6	Sequence 6, Appli
20	1590	40.1	391	6	US-08-244-951-9	Sequence 9, Appli
21	1590	40.1	391	21	US-09-790-148-23	Sequence 23, Appli
22	1590	40.1	391	21	US-09-790-148A-23	Sequence 23, Appli
23	1586	40.0	351	27	US-60-087-557-7	Sequence 7, Appli
24	1552	39.1	374	26	US-10-205-331-108	Sequence 108, App
25	1253.5	31.6	365	21	US-09-791-537-7697	Sequence 7697, Ap
26	1119	28.2	239	21	US-09-760-469-1123	Sequence 1123, Ap
27	1119	28.2	239	26	US-10-216-583-1123	Sequence 1123, Ap
28	733	18.5	140	25	US-10-107-181-9	Sequence 9, Appli
29	621.5	15.7	1125	18	US-09-430-656-152	Sequence 152, App
30	621.5	15.7	1125	21	US-09-713-508-152	Sequence 152, App
31	621.5	15.7	1125	21	US-09-713-572-152	Sequence 152, App
32	621.5	15.7	1610	18	US-09-430-656-22	Sequence 22, Appli
33	621.5	15.7	1610	21	US-09-713-508-22	Sequence 22, Appli
34	621.5	15.7	1610	21	US-09-713-572-22	Sequence 22, Appli
35	593	15.0	112	18	US-09-432-285-1	Sequence 1, Appli
36	575	14.5	109	25	US-10-107-181-3	Sequence 3, Appli
37	561	14.1	108	25	US-10-107-181-2	Sequence 2, Appli
38	546.5	13.8	109	25	US-10-107-181-1	Sequence 1, Appli
39	546	13.8	140	25	US-10-107-181-10	Sequence 10, Appli
40	495	12.5	322	21	US-09-760-469-1124	Sequence 1124, Ap
41	495	12.5	322	26	US-10-216-583-1124	Sequence 1124, Ap
42	471	11.9	700	20	US-09-614-150-11805	Sequence 11805, A
43	471	11.9	700	27	US-60-167-217-11819	Sequence 11819, A
44	471	11.9	700	27	US-60-171-625-231	Sequence 231, App
45	471	11.9	700	27	US-60-173-464-9576	Sequence 9576, App

ALIGNMENTS

RESULT 1
PCT-US02-22032-5
; Sequence 5, Application PC/TUS0222032
; GENERAL INFORMATION:
; APPLICANT: Novactyl, Inc.
; TITLE OF INVENTION: Methods and Compositions for Controlling Protein Assembly or A
; CURRENT APPLICATION NUMBER: 42108/34520
; CURRENT FILING DATE: 2002-07-11
; PRIOR APPLICATION NUMBER: 09/904,987
; PRIOR FILING DATE: 12 July 2001
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 758
; TYPE: PRT
; ORGANISM: homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NCBI ENTREZ / NM_016835
; DATABASE ENTRY DATE: 2001-02-13
; RELEVANT RESIDUES: (1)..(758)
PCT-US02-22032-5

Query Match 100.0%; Score 3965; DB 1; Length 758;
Best Local Similarity 100.0%; Pred. No. 8.6e-210;
Matches 758; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MAEPQFEVMEHAGTYGLGDRKDGQGYTHQDQEGTDAGLKESPLQPTDGSSEPG 60
|||||

Db 1 MAEPQREFEVMEDHAGTYGLGDRKQGGYTMHQDQEGDTAGLKESTPLQTTEDGSEEPG 60
Qy 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAQPHTEIPGTTAEAGIGDTPSLEDEAAG 120
Db 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAQPHTEIPGTTAEAGIGDTPSLEDEAAG 120
Qy 121 HVTQEPESGVVQEGFLREPGLSHQLSMGMPGAPLLPEGPREATROPSTGTPEDTEG 180
Db 121 HVTQEPESGVVQEGFLREPGLSHQLSMGMPGAPLLPEGPREATROPSTGTPEDTEG 180
Qy 181 GRHAPELLKHQLGLDGLHQBEGPPLKAGGKERPGSKKEEVEDRDVDDESSPODSSPKASPA 240
Db 181 GRHAPELLKHQLGLDGLHQBEGPPLKAGGKERPGSKKEEVEDRDVDDESSPODSSPKASPA 240
Qy 241 ODGRPPQTAAREATSIPIGFAEAGIPLPVDLFSKVSTEIPASEPDGSPVGRAGQDAPLE 300
Db 241 ODGRPPQTAAREATSIPIGFAEAGIPLPVDLFSKVSTEIPASEPDGSPVGRAGQDAPLE 300
Qy 301 FTFHVEITPNVQKEQAHEEHLGRAAFPGAGPEARGPSLGEDTKEADLPEPSEKQPA 360
Db 301 FTFHVEITPNVQKEQAHEEHLGRAAFPGAGPEARGPSLGEDTKEADLPEPSEKQPA 360
Qy 361 AAPRGKPVSRVQPKARWVSKSDGTGSDDKKAKTSTRSSAKTLKNRCLSPKLPPTGSS 420
Db 361 AAPRGKPVSRVQPKARWVSKSDGTGSDDKKAKTSTRSSAKTLKNRCLSPKLPPTGSS 420
Qy 421 DPLIOPSSPAVCPEPPSPKHVSSVTSRTGSSGAKEMKLGADGKTKIATPRGAAPPGQK 480
Db 421 DPLIOPSSPAVCPEPPSPKHVSSVTSRTGSSGAKEMKLGADGKTKIATPRGAAPPGQK 480
Qy 481 GOANATRIAPKTPAPKTPSSGEPKSGDRSGYSSPGSPGTPGSRSTRPSLPTPTREP 540
Db 481 GOANATRIAPKTPAPKTPSSGEPKSGDRSGYSSPGSPGTPGSRSTRPSLPTPTREP 540
Qy 541 KKVAVRTPPKSPSSAKSLQTPAPVMPDLKNVSKSTIGSTENLKHQPGGKQVQIINKLD 600
Db 541 KKVAVRTPPKSPSSAKSLQTPAPVMPDLKNVSKSTIGSTENLKHQPGGKQVQIINKLD 600
Qy 601 LSNVQSKCGSKDNKIHVPGGGSVQIVYKPVDSLKVTSKCGSLGNIHHKPGGGQVEVKSEK 660
Db 601 LSNVQSKCGSKDNKIHVPGGGSVQIVYKPVDSLKVTSKCGSLGNIHHKPGGGQVEVKSEK 660
Qy 661 LDFKDRVQSKIGSLDNITHVPGGKNKIETHKLTFRENAKAKTDHGAEIVYKSPVVSQDT 720
Db 661 LDFKDRVQSKIGSLDNITHVPGGKNKIETHKLTFRENAKAKTDHGAEIVYKSPVVSQDT 720
Qy 721 SPRHLNSVSTGSDMDVDSPLATLADEVASLAKQGL 758
Db 721 SPRHLNSVSTGSDMDVDSPLATLADEVASLAKQGL 758

RESULT 2

US-09-904-987-5
; Sequence 5, Application US/09904987
; GENERAL INFORMATION:
; APPLICANT: Novactyl, Inc.
; TITLE OF INVENTION: Methods and Compositions for Controlling Pathological and Prepath
; TITLE OF INVENTION: Protein Assembly or Aggregation
; FILE REFERENCE: 42108/26146
; CURRENT APPLICATION NUMBER: US/09/904,987
; CURRENT FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 758
; TYPE: PRT
; ORGANISM: homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NCBI ENTREZ / NM_016835
; DATABASE ENTRY DATE: 2001-02-13
; RELEVANT RESIDUES: (1)..(758)
US-09-904-987-5

Query Match 100.0%; Score 3965; DB 23; Length 758;
Best Local Similarity 100.0%; Pred. No. 8.6e-210;
Matches 758; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MAEPQREFEVMEDHAGTYGLGDRKQGGYTMHQDQEGDTAGLKESTPLQTTEDGSEEPG 60
Db 1 MAEPQREFEVMEDHAGTYGLGDRKQGGYTMHQDQEGDTAGLKESTPLQTTEDGSEEPG 60
Qy 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAQPHTEIPGTTAEAGIGDTPSLEDEAAG 120
Db 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAQPHTEIPGTTAEAGIGDTPSLEDEAAG 120
Qy 121 HVTQEPESGVVQEGFLREPGLSHQLSMGMPGAPLLPEGPREATROPSTGTPEDTEG 180
Db 121 HVTQEPESGVVQEGFLREPGLSHQLSMGMPGAPLLPEGPREATROPSTGTPEDTEG 180
Qy 181 GRHAPELLKHQLGLDGLHQBEGPPLKAGGKERPGSKKEEVEDRDVDDESSPODSSPKASPA 240
Db 181 GRHAPELLKHQLGLDGLHQBEGPPLKAGGKERPGSKKEEVEDRDVDDESSPODSSPKASPA 240
Qy 241 ODGRPPQTAAREATSIPIGFAEAGIPLPVDLFSKVSTEIPASEPDGSPVGRAGQDAPLE 300
Db 241 ODGRPPQTAAREATSIPIGFAEAGIPLPVDLFSKVSTEIPASEPDGSPVGRAGQDAPLE 300
Qy 301 FTFHVEITPNVQKEQAHEEHLGRAAFPGAGPEARGPSLGEDTKEADLPEPSEKQPA 360
Db 301 FTFHVEITPNVQKEQAHEEHLGRAAFPGAGPEARGPSLGEDTKEADLPEPSEKQPA 360
Qy 361 AAPRGKPVSRVQPKARWVSKSDGTGSDDKKAKTSTRSSAKTLKNRCLSPKLPPTGSS 420
Db 361 AAPRGKPVSRVQPKARWVSKSDGTGSDDKKAKTSTRSSAKTLKNRCLSPKLPPTGSS 420
Qy 421 DPLIOPSSPAVCPEPPSPKHVSSVTSRTGSSGAKEMKLGADGKTKIATPRGAAPPGQK 480
Db 421 DPLIOPSSPAVCPEPPSPKHVSSVTSRTGSSGAKEMKLGADGKTKIATPRGAAPPGQK 480
Qy 481 GOANATRIAPKTPAPKTPSSGEPKSGDRSGYSSPGSPGTPGSRSTRPSLPTPTREP 540
Db 481 GOANATRIAPKTPAPKTPSSGEPKSGDRSGYSSPGSPGTPGSRSTRPSLPTPTREP 540
Qy 541 KKVAVRTPPKSPSSAKSLQTPAPVMPDLKNVSKSTIGSTENLKHQPGGKQVQIINKLD 600
Db 541 KKVAVRTPPKSPSSAKSLQTPAPVMPDLKNVSKSTIGSTENLKHQPGGKQVQIINKLD 600
Qy 601 LSNVQSKCGSKDNKIHVPGGGSVQIVYKPVDSLKVTSKCGSLGNIHHKPGGGQVEVKSEK 660
Db 601 LSNVQSKCGSKDNKIHVPGGGSVQIVYKPVDSLKVTSKCGSLGNIHHKPGGGQVEVKSEK 660
Qy 661 LDFKDRVQSKIGSLDNITHVPGGKNKIETHKLTFRENAKAKTDHGAEIVYKSPVVSQDT 720
Db 661 LDFKDRVQSKIGSLDNITHVPGGKNKIETHKLTFRENAKAKTDHGAEIVYKSPVVSQDT 720
Qy 721 SPRHLNSVSTGSDMDVDSPLATLADEVASLAKQGL 758
Db 721 SPRHLNSVSTGSDMDVDSPLATLADEVASLAKQGL 758

RESULT 3

US-60-406-385-228
; Sequence 228, Application US/60406385
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION OF GENES FOR PREDICTING ACTIVITY OF COMPOUNDS T
; TITLE OF INVENTION: INTERACT WITH SRC TYROSINE KINASES AND THE SRC TYROSINE KINAS
; TITLE OF INVENTION: CANCER CELLS
; FILE REFERENCE: D0273 PSP
; CURRENT APPLICATION NUMBER: US/60/406,385
; CURRENT FILING DATE: 2002-08-27
; NUMBER OF SEQ ID NOS: 516
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 228
; LENGTH: 758
; TYPE: PRT

ORGANISM: Homo sapiens
US-60-406-385-228

```

Query Match      100.0%; Score 3965; DB 27; Length 758;
Best Local Similarity 100.0%; Pred. No. 8.6e-210;
Matches 758; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAEPQEFVEVMEHAGTYGLGDRKDGQYTMHODQEGDTDAGLKESPLQTTPTEDGSEEPG 60
   |||||||
Db 1 MAEPQEFVEVMEHAGTYGLGDRKDGQYTMHODQEGDTDAGLKESPLQTTPTEDGSEEPG 60

QY 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAQAHPTEIPGTTAEAGIGDTPSLEDEAAG 120
   |||||||
Db 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAQAHPTEIPGTTAEAGIGDTPSLEDEAAG 120

QY 121 HVTQEPESGKVQVEGFLREPGLSHQMSGMPGAPLPPGEPREATRQPSGTGPEDETEG 180
   |||||||
Db 121 HVTQEPESGKVQVEGFLREPGLSHQMSGMPGAPLPPGEPREATRQPSGTGPEDETEG 180

QY 181 GRHAPELLKHOLLGDLHOEGPPLKAGGKERPGSKKEEVEDRDVDESSPDSPSKASPA 240
   |||||||
Db 181 GRHAPELLKHOLLGDLHOEGPPLKAGGKERPGSKKEEVEDRDVDESSPDSPSKASPA 240

QY 241 QDGRPPQTAAREATSIPGPPAEGAIPLVDFLSKVSTEIPASEPDGSPVGRAGQDAPLE 300
   |||||||
Db 241 QDGRPPQTAAREATSIPGPPAEGAIPLVDFLSKVSTEIPASEPDGSPVGRAGQDAPLE 300

QY 301 FTFHVETPNVQEQAHSEHLGRAAFPGAPGEGPARGPSLGEDTKADLPPESEKQPA 360
   |||||||
Db 301 FTFHVETPNVQEQAHSEHLGRAAFPGAPGEGPARGPSLGEDTKADLPPESEKQPA 360

QY 361 AAPRGKPVSRVLPOLKARMVSKSKDGTGSDDKAKTSTRSSAKTLKNRCLSPKLPPTGSS 420
   |||||||
Db 361 AAPRGKPVSRVLPOLKARMVSKSKDGTGSDDKAKTSTRSSAKTLKNRCLSPKLPPTGSS 420

QY 421 DPLIQSPSPAVCPPEPPSPKHSVSVTSRTGSSGAKEMKLGADGKTKIATPRGAAPPQOK 480
   |||||||
Db 421 DPLIQSPSPAVCPPEPPSPKHSVSVTSRTGSSGAKEMKLGADGKTKIATPRGAAPPQOK 480

QY 481 GOANATRIPAKTPAPKTPPSSGEPKSGDRSGYSSPGSPGTPGSRSPSLPTPTPTREP 540
   |||||||
Db 481 GOANATRIPAKTPAPKTPPSSGEPKSGDRSGYSSPGSPGTPGSRSPSLPTPTPTREP 540

QY 541 KKVAVVRTPPKSPSSAKSLQATAPVMPDLKNVKSIGSTENLKHOPGGGKQVQIINKKLD 600
   |||||||
Db 541 KKVAVVRTPPKSPSSAKSLQATAPVMPDLKNVKSIGSTENLKHOPGGGKQVQIINKKLD 600

QY 601 LSNVQSKGSKDNKIHVPGGGSGVQIYKPVDSLKVTSCGSLGNIHHKPGGQGVVEKSEK 660
   |||||||
Db 601 LSNVQSKGSKDNKIHVPGGGSGVQIYKPVDSLKVTSCGSLGNIHHKPGGQGVVEKSEK 660

QY 661 LDPKDRVQSKIGSLDNITHVPGGNGKKIETHKLTFRENAKAKTDHGAETVYKSPVVSQGT 720
   |||||||
Db 661 LDPKDRVQSKIGSLDNITHVPGGNGKKIETHKLTFRENAKAKTDHGAETVYKSPVVSQGT 720

QY 721 SPRHLNVSTSGSIDMVDSPQLATLADEVASLAKOGL 758
   |||||||
Db 721 SPRHLNVSTSGSIDMVDSPQLATLADEVASLAKOGL 758

```

RESULT 4

```

US-08-477-648-1
: Sequence 1. Application US/08477648
: GENERAL INFORMATION:
: APPLICANT: Mandelkow, Eva-Maria
: APPLICANT: Mandelkow, Eckhard
: APPLICANT: Lichtenberg-Kraag, Birgit
: APPLICANT: Biernat, Jacek
: APPLICANT: Drewes, Gerard
: APPLICANT: Steiner, Barbara
: TITLE OF INVENTION: Novel Tools For The Diagnosis And
: TITLE OF INVENTION: Treatment Of Alzheimer's Disease
: NUMBER OF SEQUENCES: 1

```

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Tape
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,648
FILING DATE: June 7, 1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Joseph A. Williams, Jr.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 28384/32777
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-484-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 441 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-477-648-1

```

```

Query Match      53.4%; Score 2116.5; DB 8; Length 441;
Best Local Similarity 58.2%; Pred. No. 3.2e-108;
Matches 441; Conservative 0; Mismatches 0; Indels 317; Gaps 2;

QY 1 MAEPQEFVEVMEHAGTYGLGDRKDGQYTMHODQEGDTDAGLKESPLQTTPTEDGSEEPG 60
   |||||||
Db 1 MAEPQEFVEVMEHAGTYGLGDRKDGQYTMHODQEGDTDAGLKESPLQTTPTEDGSEEPG 60

QY 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAQAHPTEIPGTTAEAGIGDTPSLEDEAAG 120
   |||||||
Db 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAQAHPTEIPGTTAEAGIGDTPSLEDEAAG 120

QY 121 HVTQEPESGKVQVEGFLREPGLSHQMSGMPGAPLPPGEPREATRQPSGTGPEDETEG 180
   |||||
Db 121 HVTQ----- 124

QY 181 GRHAPELLKHOLLGDLHOEGPPLKAGGKERPGSKKEEVEDRDVDESSPDSPSKASPA 240
   |||||||
Db 181 GRHAPELLKHOLLGDLHOEGPPLKAGGKERPGSKKEEVEDRDVDESSPDSPSKASPA 240

QY 241 QDGRPPQTAAREATSIPGPPAEGAIPLVDFLSKVSTEIPASEPDGSPVGRAGQDAPLE 300
   |||||||
Db 241 QDGRPPQTAAREATSIPGPPAEGAIPLVDFLSKVSTEIPASEPDGSPVGRAGQDAPLE 300

QY 301 FTFHVETPNVQEQAHSEHLGRAAFPGAPGEGPARGPSLGEDTKADLPPESEKQPA 360
   |||||||
Db 301 FTFHVETPNVQEQAHSEHLGRAAFPGAPGEGPARGPSLGEDTKADLPPESEKQPA 360

QY 361 AAPRGKPVSRVLPOLKARMVSKSKDGTGSDDKAKTSTRSSAKTLKNRCLSPKLPPTGSS 420
   |||||||
Db 361 AAPRGKPVSRVLPOLKARMVSKSKDGTGSDDKAKTSTRSSAKTLKNRCLSPKLPPTGSS 420

QY 421 DPLIQSPSPAVCPPEPPSPKHSVSVTSRTGSSGAKEMKLGADGKTKIATPRGAAPPQOK 480
   |||||||
Db 421 DPLIQSPSPAVCPPEPPSPKHSVSVTSRTGSSGAKEMKLGADGKTKIATPRGAAPPQOK 480

QY 481 GOANATRIPAKTPAPKTPPSSGEPKSGDRSGYSSPGSPGTPGSRSPSLPTPTPTREP 540
   |||||||
Db 481 GOANATRIPAKTPAPKTPPSSGEPKSGDRSGYSSPGSPGTPGSRSPSLPTPTPTREP 540

QY 541 KKVAVVRTPPKSPSSAKSLQATAPVMPDLKNVKSIGSTENLKHOPGGGKQVQIINKKLD 600
   |||||||
Db 541 KKVAVVRTPPKSPSSAKSLQATAPVMPDLKNVKSIGSTENLKHOPGGGKQVQIINKKLD 600

QY 601 LSNVQSKGSKDNKIHVPGGGSGVQIYKPVDSLKVTSCGSLGNIHHKPGGQGVVEKSEK 660
   |||||||
Db 601 LSNVQSKGSKDNKIHVPGGGSGVQIYKPVDSLKVTSCGSLGNIHHKPGGQGVVEKSEK 660

QY 661 LDPKDRVQSKIGSLDNITHVPGGNGKKIETHKLTFRENAKAKTDHGAETVYKSPVVSQGT 720
   |||||||
Db 661 LDPKDRVQSKIGSLDNITHVPGGNGKKIETHKLTFRENAKAKTDHGAETVYKSPVVSQGT 720

QY 721 SPRHLNVSTSGSIDMVDSPQLATLADEVASLAKOGL 758
   |||||||
Db 721 SPRHLNVSTSGSIDMVDSPQLATLADEVASLAKOGL 758

```

QY 541 KKVAVVTPPKSPSSAKSLRLQAPVMPDLKNVSKIGSTENLKHQPGGKVQIINKKLD 600
 DB 224 KKVAVVTPPKSPSSAKSLRLQAPVMPDLKNVSKIGSTENLKHQPGGKVQIINKKLD 283
 QY 601 LSNVQSKGSKDNKIHVPGGSGVQIVYKPDLSKVTSCGSLGNHKKHKKPGGQGVKSEK 660
 DB 284 LSNVQSKGSKDNKIHVPGGSGVQIVYKPDLSKVTSCGSLGNHKKHKKPGGQGVKSEK 343
 QY 661 LDFKDRVQSKIGSLDNITHVPGGSKKTIETHKLTIFRENAKAKTDHGAIEIVYKSPVVS GDT 720
 DB 344 LDFKDRVQSKIGSLDNITHVPGGSKKTIETHKLTIFRENAKAKTDHGAIEIVYKSPVVS GDT 403
 QY 721 SPRHLNSVSTGSDIMVDSPLATLADEVASLAKOGL 758
 DB 404 SPRHLNSVSTGSDIMVDSPLATLADEVASLAKOGL 441

RESULT 5

US-08-484-674-1
 ; Sequence 1, Application US/08484674
 ; GENERAL INFORMATION:
 ; APPLICANT: Mandelkow, Eva-Maria
 ; APPLICANT: Mandelkow, Eckhard
 ; APPLICANT: Lichtenberg-Kraag, Birgit
 ; APPLICANT: Biernat, Jacek
 ; APPLICANT: Drewes, Gerard
 ; APPLICANT: Steiner, Barbara
 ; TITLE OF INVENTION: Novel Tools For The Diagnosis And
 ; TITLE OF INVENTION: Treatment Of Alzheimer's Disease
 ; NUMBER OF SEQUENCES: 1
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 ; STREET: 233 South Wacker Drive, 6300 Sears Tower
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: United States of America
 ; ZIP: 60606-6402
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Tape
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/484,674
 ; FILING DATE: June 7, 1995
 ; CLASSIFICATION: 424
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Joseph A. Williams, Jr.
 ; REGISTRATION NUMBER: 38,659
 ; REFERENCE/DOCKET NUMBER: 28384/32778
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 312-474-6300
 ; TELEFAX: 312-484-0448
 ; TELEX: 25-3856
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 441 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-484-674-1

Query Match 53.4%; Score 2116.5; DB 8; Length 441;
 Best Local Similarity 58.2%; Pred. No. 3 2e-108;
 Matches 441; Conservative 0; Mismatches 0; Indels 317; Gaps 2;
 QY 1 MAEPROFEFVEMDHAGTYGLGDRKDGQYTMHQQDQDGTDLAKESPLQTPTEDSSEEPG 60
 DB 1 MAEPROFEFVEMDHAGTYGLGDRKDGQYTMHQQDQDGTDLAKESPLQTPTEDSSEEPG 60
 QY 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAQPHTEIPGTTAEAGIGDTPSLEDEAAG 120

DB 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAQPHTEIPGTTAEAGIGDTPSLEDEAAG 120
 QY 121 HVTQEPESGKVVQEGFLREPQPGGLSHQLMSGMPGAPLLPEGPREATRQPSGTGPEDETEG 180
 DB 121 HVTQ----- 124
 QY 181 GRHAPELLKHOLLGDLHQEGPPLKAGGKERPGSKKEEVEDROVDSESSPODSPSKASPA 240
 DB 125 ----- 124
 QY 241 QDGRPPQTAAREATSI PCFPAEGAIPLPVDFLSKVSTEIPASEPDGFSVGRAKGQDAPLE 300
 DB 125 ----- 124
 QY 301 FTFHVEITPNVQKEQAHSEHGLGRAFPAGPEGGEARGPSLGEDTKREADLPEPSEKQPA 360
 DB 125 ----- 124
 QY 361 AAPRGKPVSRYPQLKARVSKSKDGTGSDKKAKTSTRSSAKTLKNRCLSPKLPTPGSS 420
 DB 125 -----ARMVSKSKDGTGSDDKA----- 142
 QY 421 DPLIQSSPAVCPPEPPSPKHSVSTSRSTGSSGAKEMKLGADGKTKIATPRGAAPPGQK 480
 DB 143 -----KGADGKTKIATPRGAAPPGQK 163
 QY 481 QANATRIAPAKTPPAPKTPPSSGEPKSGDRSGYSSPGSPGTPGSRSTPSLPTPTPREP 540
 DB 164 QANATRIAPAKTPPAPKTPPSSGEPKSGDRSGYSSPGSPGTPGSRSTPSLPTPTPREP 223
 QY 541 KKVAVVTPPKSPSSAKSLRLQAPVMPDLKNVSKIGSTENLKHQPGGKVQIINKKLD 600
 DB 224 KKVAVVTPPKSPSSAKSLRLQAPVMPDLKNVSKIGSTENLKHQPGGKVQIINKKLD 283
 QY 601 LSNVQSKGSKDNKIHVPGGSGVQIVYKPDLSKVTSCGSLGNHKKHKKPGGQGVKSEK 660
 DB 284 LSNVQSKGSKDNKIHVPGGSGVQIVYKPDLSKVTSCGSLGNHKKHKKPGGQGVKSEK 343
 QY 661 LDFKDRVQSKIGSLDNITHVPGGSKKTIETHKLTIFRENAKAKTDHGAIEIVYKSPVVS GDT 720
 DB 344 LDFKDRVQSKIGSLDNITHVPGGSKKTIETHKLTIFRENAKAKTDHGAIEIVYKSPVVS GDT 403
 QY 721 SPRHLNSVSTGSDIMVDSPLATLADEVASLAKOGL 758
 DB 404 SPRHLNSVSTGSDIMVDSPLATLADEVASLAKOGL 441

RESULT 6

US-09-035-708-1
 ; Sequence 1, Application US/09035708
 ; GENERAL INFORMATION:
 ; APPLICANT: University of Cincinnati
 ; TITLE OF INVENTION: METHOD OF DETECTING AXONAL DAMAGE, ASSOCIATED DISEASE
 ; TITLE OF INVENTION: STATES, AND RELATED MONOCLONAL ANTIBODIES AND PROTEIN
 ; TITLE OF INVENTION: CONTROLS THEREFOR
 ; FILE REFERENCE: 1259-064
 ; CURRENT APPLICATION NUMBER: US/09/035,708
 ; CURRENT FILING DATE: 1999-03-05
 ; NUMBER OF SEQ ID NOS: 1
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 441
 ; TYPE: PRT
 ; ORGANISM: human clone ht4u40 isoform reduced
 ; PUBLICATION INFORMATION:
 ; AUTHORS: Goedert, M.
 ; TITLE: Multiple Isoforms of Human Microtubule-Associated
 ; TITLE: Protein Tau: Sequences and Localization in
 ; TITLE: Neurofibrillary Tangles of Alzheimer's Disease
 ; JOURNAL: Neuron
 ; VOLUME: 3
 ; PAGES: 519-526
 ; DATE: OCT-1989

US-09-035-708-1

```
Query Match      53.4%; Score 2116.5; DB 14; Length 441;
Best Local Similarity 58.2%; Pred. No. 3.2e-108;
Matches 441; Conservative 0; Mismatches 0; Indels 317; Gaps 2;

QY 1 MAEPQEFVEFVEMDHAGTYGLGDRKDGQGYTMHQDQEGDTAGLKESPLQTPTEGSEEPG 60
   |||||
Db 1 MAEPQEFVEFVEMDHAGTYGLGDRKDGQGYTMHQDQEGDTAGLKESPLQTPTEGSEEPG 60

QY 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAAPHTPEIGTTAEAGTGTDEAAG 120
   |||||
Db 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAAPHTPEIGTTAEAGTGTDEAAG 120

QY 121 HVTQEPESGVVQEGFLREPGLSHQMSGMPGAPLLPEGPREATRQPSGTGPDTEG 180
   |||||
Db 121 HVTQ-----

QY 181 GRHAPELLKHOLLGDLHQEGPPLKGAGKERPGSKEEVEDRDVDESSPDSPSKASPA 240
   |||||
Db 125 -----

QY 241 QDGRPPQTAAREATSI PGPAEGAIPLVDFLSKVSTEIPASEPDGPGSVGRAKQDAPLE 300
   |||||
Db 125 -----

QY 301 FTFHVEITPNVQKQAHSEHLGRAAPFGAPGEGPEARGPSLGEDTKREADLPEPSEKQPA 360
   |||||
Db 125 -----

QY 361 AAPRGKPVSRVQPKARVMSKSDGTGSDDKKAKTSTRSSAKTLKNRCLSPKLPTPGSS 420
   |||||
Db 125 -----ARMVSKSDGTGSDDKKA-----

QY 421 DPLIQSSPAVCEPPSPKHSVTSRTGSSGAKEMKLGADGKTATPRGAAPPQK 480
   |||||
Db 143 -----KGADGKTATPRGAAPPQK 163

QY 481 QGANATRIPAKTPAPKTPPSSGEPKSGDRSGYSSPGSGTPGSRRTPSLPTPTREP 540
   |||||
Db 164 QGANATRIPAKTPAPKTPPSSGEPKSGDRSGYSSPGSGTPGSRRTPSLPTPTREP 223

QY 541 KKVAVVTRTPKSPSSAKSLQTAAPVMPDLKNVSKIGSTENLKHQPGGKQVQIINKKLD 600
   |||||
Db 224 KKVAVVTRTPKSPSSAKSLQTAAPVMPDLKNVSKIGSTENLKHQPGGKQVQIINKKLD 283

QY 601 LSNVQSKCGSKDNIKHVPGGGSQVIVYKPVDSLKVTSCGSLGNTHHHPGGQGVKSEK 660
   |||||
Db 284 LSNVQSKCGSKDNIKHVPGGGSQVIVYKPVDSLKVTSCGSLGNTHHHPGGQGVKSEK 343

QY 661 LDFKDRVQSKIGSLDNITHVPGGKNKIETHKLTFRENAKAKTDHGAEIVYKSPVVSGDT 720
   |||||
Db 344 LDFKDRVQSKIGSLDNITHVPGGKNKIETHKLTFRENAKAKTDHGAEIVYKSPVVSGDT 403

QY 721 SPRHLSNVSTGSDIMVDSQPLATLADEVASLAKQGL 758
   |||||
Db 404 SPRHLSNVSTGSDIMVDSQPLATLADEVASLAKQGL 441

RESULT 7
US-09-035-708A-1
; Sequence 1, Application US/09035708A
; GENERAL INFORMATION:
; APPLICANT: University of Cincinnati
; TITLE OF INVENTION: METHOD OF DETECTING AXONAL DAMAGE, ASSOCIATED DISEASE
; TITLE OF INVENTION: STATES AND RELATED MONOCLONAL ANTIBODIES AND PROTEIN
; TITLE OF INVENTION: CONTROLS THEREFOR
; FILE REFERENCE: 1259-064
; CURRENT APPLICATION NUMBER: US/09/035.708A
; CURRENT FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
```

; LENGTH: 441

; TYPE: PRT

; ORGANISM: human clone ht40 isoform reduced

; PUBLICATION INFORMATION:

; AUTHORS: Goedert, M.

; TITLE: Multiple Isoforms of Human Microtubule-Associated Protein Tau:

; TITLE: Sequences and Localization in Neurofibrillary Tangles of Alzheimer's

; TITLE: Disease

; JOURNAL: Neuron

; VOLUME: 3

; PAGES: 519-526

; DATE: OCT-1989

US-09-035-708A-1

Query Match

Best Local Similarity 58.2%; Pred. No. 3.2e-108;

Matches 441; Conservative 0; Mismatches 0; Indels 317; Gaps 2;

QY 1 MAEPQEFVEFVEMDHAGTYGLGDRKDGQGYTMHQDQEGDTAGLKESPLQTPTEGSEEPG 60

Db 1 MAEPQEFVEFVEMDHAGTYGLGDRKDGQGYTMHQDQEGDTAGLKESPLQTPTEGSEEPG 60

QY 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAAPHTPEIGTTAEAGTGTDEAAG 120

Db 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAAPHTPEIGTTAEAGTGTDEAAG 120

QY 121 HVTQEPESGVVQEGFLREPGLSHQMSGMPGAPLLPEGPREATRQPSGTGPDTEG 180

Db 121 HVTQ-----

QY 181 GRHAPELLKHOLLGDLHQEGPPLKGAGKERPGSKEEVEDRDVDESSPDSPSKASPA 240

Db 125 -----

QY 241 QDGRPPQTAAREATSI PGPAEGAIPLVDFLSKVSTEIPASEPDGPGSVGRAKQDAPLE 300

Db 125 -----

QY 301 FTFHVEITPNVQKQAHSEHLGRAAPFGAPGEGPEARGPSLGEDTKREADLPEPSEKQPA 360

Db 125 -----

QY 361 AAPRGKPVSRVQPKARVMSKSDGTGSDDKKAKTSTRSSAKTLKNRCLSPKLPTPGSS 420

Db 125 -----ARMVSKSDGTGSDDKKA-----

QY 421 DPLIQSSPAVCEPPSPKHSVTSRTGSSGAKEMKLGADGKTATPRGAAPPQK 480

Db 143 -----KGADGKTATPRGAAPPQK 163

QY 481 QGANATRIPAKTPAPKTPPSSGEPKSGDRSGYSSPGSGTPGSRRTPSLPTPTREP 540

Db 164 QGANATRIPAKTPAPKTPPSSGEPKSGDRSGYSSPGSGTPGSRRTPSLPTPTREP 223

QY 541 KKVAVVTRTPKSPSSAKSLQTAAPVMPDLKNVSKIGSTENLKHQPGGKQVQIINKKLD 600

Db 224 KKVAVVTRTPKSPSSAKSLQTAAPVMPDLKNVSKIGSTENLKHQPGGKQVQIINKKLD 283

QY 601 LSNVQSKCGSKDNIKHVPGGGSQVIVYKPVDSLKVTSCGSLGNTHHHPGGQGVKSEK 660

Db 284 LSNVQSKCGSKDNIKHVPGGGSQVIVYKPVDSLKVTSCGSLGNTHHHPGGQGVKSEK 343

QY 661 LDFKDRVQSKIGSLDNITHVPGGKNKIETHKLTFRENAKAKTDHGAEIVYKSPVVSGDT 720

Db 344 LDFKDRVQSKIGSLDNITHVPGGKNKIETHKLTFRENAKAKTDHGAEIVYKSPVVSGDT 403

QY 721 SPRHLSNVSTGSDIMVDSQPLATLADEVASLAKQGL 758

Db 404 SPRHLSNVSTGSDIMVDSQPLATLADEVASLAKQGL 441

RESULT 8

US-09-986-347-71

```

1  ; Sequence 71, Application US/09386347
2  ;
3  ; GENERAL INFORMATION:
4  ;
5  ; APPLICANT: Daniel M. Michaelson et al.
6  ;
7  ; TITLE OF INVENTION: DEVICES AND SUBSTANCES, METHODS AND
8  ;
9  ; TITLE OF INVENTION: DEVICES USING SAME FOR DIAGNOSING AND
10 ;
11 ; TITLE OF INVENTION: TREATING NEURODEGENERATIVE DISORDERS
12 ;
13 ; NUMBER OF SEQUENCES: 71
14 ;
15 ; CORRESPONDENCE ADDRESS:
16 ;
17 ; ADDRESS: Mark M. Friedman c/o Anthony Castorina
18 ;
19 ; STREET: 2001 Jefferson Davis Highway, Suite 207
20 ;
21 ; CITY: Arlington
22 ;
23 ; STATE: Virginia
24 ;
25 ; COUNTRY: United States of America
26 ;
27 ; ZIP: 22202
28 ;
29 ; COMPUTER READABLE FORM:
30 ;
31 ; MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk
32 ;
33 ; COMPUTER: Twinhead* Slimnote-890TX
34 ;
35 ; OPERATING SYSTEM: MS DOS version 6.2,
36 ;
37 ; OPERATING SYSTEM: Windows version 3.11
38 ;
39 ; SOFTWARE: Word for Windows version 2.0
40 ;
41 ; SOFTWARE: converted to an ASCII file
42 ;
43 ; CURRENT APPLICATION DATA:
44 ;
45 ; APPLICATION NUMBER: US/09/386,347
46 ;
47 ; FILING DATE:
48 ;
49 ; CLASSIFICATION:
50 ;
51 ; PRIOR APPLICATION DATA:
52 ;
53 ; APPLICATION NUMBER:
54 ;
55 ; FILING DATE:
56 ;
57 ; ATTORNEY/AGENT INFORMATION:
58 ;
59 ; NAME: Friedmam, Mark M.
60 ;
61 ; REGISTRATION NUMBER: 33,883
62 ;
63 ; REFERENCE/DOCKET NUMBER: 1710/1
64 ;
65 ; TELECOMMUNICATION INFORMATION:
66 ;
67 ; TELEPHONE: 972-3-5625553
68 ;
69 ; TELEFAX: 972-3-5625554
70 ;
71 ; TELEX:
72 ;
73 ; INFORMATION FOR SEQ ID NO: 71:
74 ;
75 ; SEQUENCE CHARACTERISTICS:
76 ;
77 ; LENGTH: 441
78 ;
79 ; TYPE: amino acid
80 ;
81 ; STRANDEDNESS: single
82 ;
83 ; TOPOLOGY: linear
84 ;
85 ;
86 ;
87 ;
88 ;
89 ;
90 ;
91 ;
92 ;
93 ;
94 ;
95 ;
96 ;
97 ;
98 ;
99 ;
100 ;
101 ;
102 ;
103 ;
104 ;
105 ;
106 ;
107 ;
108 ;
109 ;
110 ;
111 ;
112 ;
113 ;
114 ;
115 ;
116 ;
117 ;
118 ;
119 ;
120 ;
121 ;
122 ;
123 ;
124 ;
125 ;
126 ;
127 ;
128 ;
129 ;
130 ;
131 ;
132 ;
133 ;
134 ;
135 ;
136 ;
137 ;
138 ;
139 ;
140 ;
141 ;
142 ;
143 ;
144 ;
145 ;
146 ;
147 ;
148 ;
149 ;
150 ;
151 ;
152 ;
153 ;
154 ;
155 ;
156 ;
157 ;
158 ;
159 ;
160 ;
161 ;
162 ;
163 ;
164 ;
165 ;
166 ;
167 ;
168 ;
169 ;
170 ;
171 ;
172 ;
173 ;
174 ;
175 ;
176 ;
177 ;
178 ;
179 ;
180 ;
181 ;
182 ;
183 ;
184 ;
185 ;
186 ;
187 ;
188 ;
189 ;
190 ;
191 ;
192 ;
193 ;
194 ;
195 ;
196 ;
197 ;
198 ;
199 ;
200 ;
201 ;
202 ;
203 ;
204 ;
205 ;
206 ;
207 ;
208 ;
209 ;
210 ;
211 ;
212 ;
213 ;
214 ;
215 ;
216 ;
217 ;
218 ;
219 ;
220 ;
221 ;
222 ;
223 ;
224 ;
225 ;
226 ;
227 ;
228 ;
229 ;
230 ;
231 ;
232 ;
233 ;
234 ;
235 ;
236 ;
237 ;
238 ;
239 ;
240 ;
241 ;
242 ;
243 ;
244 ;
245 ;
246 ;
247 ;
248 ;
249 ;
250 ;
251 ;
252 ;
253 ;
254 ;
255 ;
256 ;
257 ;
258 ;
259 ;
260 ;
261 ;
262 ;
263 ;
264 ;
265 ;
266 ;
267 ;
268 ;
269 ;
270 ;
271 ;
272 ;
273 ;
274 ;
275 ;
276 ;
277 ;
278 ;
279 ;
280 ;
281 ;
282 ;
283 ;
284 ;
285 ;
286 ;
287 ;
288 ;
289 ;
290 ;
291 ;
292 ;
293 ;
294 ;
295 ;
296 ;
297 ;
298 ;
299 ;
300 ;
301 ;
302 ;
303 ;
304 ;
305 ;
306 ;
307 ;
308 ;
309 ;
310 ;
311 ;
312 ;
313 ;
314 ;
315 ;
316 ;
317 ;
318 ;
319 ;
320 ;
321 ;
322 ;
323 ;
324 ;
325 ;
326 ;
327 ;
328 ;
329 ;
330 ;
331 ;
332 ;
333 ;
334 ;
335 ;
336 ;
337 ;
338 ;
339 ;
340 ;
341 ;
342 ;
343 ;
344 ;
345 ;
346 ;
347 ;
348 ;
349 ;
350 ;
351 ;
352 ;
353 ;
354 ;
355 ;
356 ;
357 ;
358 ;
359 ;
360 ;
361 ;
362 ;
363 ;
364 ;
365 ;
366 ;
367 ;
368 ;
369 ;
370 ;
371 ;
372 ;
373 ;
374 ;
375 ;
376 ;
377 ;
378 ;
379 ;
380 ;
381 ;
382 ;
383 ;
384 ;
385 ;
386 ;
387 ;
388 ;
389 ;
390 ;
391 ;
392 ;
393 ;
394 ;
395 ;
396 ;
397 ;
398 ;
399 ;
400 ;
401 ;
402 ;
403 ;
404 ;
405 ;
406 ;
407 ;
408 ;
409 ;
410 ;
411 ;
412 ;
413 ;
414 ;
415 ;
416 ;
417 ;
418 ;
419 ;
420 ;
421 ;
422 ;
423 ;
424 ;
425 ;
426 ;
427 ;
428 ;
429 ;
430 ;
431 ;
432 ;
433 ;
434 ;
435 ;
436 ;
437 ;
438 ;
439 ;
440 ;
441 ;
442 ;
443 ;
444 ;
445 ;
446 ;
447 ;
448 ;
449 ;
450 ;
451 ;
452 ;
453 ;
454 ;
455 ;
456 ;
457 ;
458 ;
459 ;
460 ;
461 ;
462 ;
463 ;
464 ;
465 ;
466 ;
467 ;
468 ;
469 ;
470 ;
471 ;
472 ;
473 ;
474 ;
475 ;
476 ;
477 ;
478 ;
479 ;
480 ;
481 ;
482 ;
483 ;
484 ;
485 ;
486 ;
487 ;
488 ;
489 ;
490 ;
491 ;
492 ;
493 ;
494 ;
495 ;
496 ;
497 ;
498 ;
499 ;
500 ;
501 ;
502 ;
503 ;
504 ;
505 ;
506 ;
507 ;
508 ;
509 ;
510 ;
511 ;
512 ;
513 ;
514 ;
515 ;
516 ;
517 ;
518 ;
519 ;
520 ;
521 ;
522 ;
523 ;
524 ;
525 ;
526 ;
527 ;
528 ;
529 ;
530 ;
531 ;
532 ;
533 ;
534 ;
535 ;
536 ;
537 ;
538 ;
539 ;
540 ;
541 ;
542 ;
543 ;
544 ;
545 ;
546 ;
547 ;
548 ;
549 ;
550 ;
551 ;
552 ;
553 ;
554 ;
555 ;
556 ;
557 ;
558 ;
559 ;
560 ;
561 ;
562 ;
563 ;
564 ;
565 ;
566 ;
567 ;
568 ;
569 ;
570 ;
571 ;
572 ;
573 ;
574 ;
575 ;
576 ;
577 ;
578 ;
579 ;
580 ;
581 ;
582 ;
583 ;
584 ;
585 ;
586 ;
587 ;
588 ;
589 ;
590 ;
591 ;
592 ;
593 ;
594 ;
595 ;
596 ;
597 ;
598 ;
599 ;
600 ;
601 ;
602 ;
603 ;
604 ;
605 ;
606 ;
607 ;
608 ;
609 ;
610 ;
611 ;
612 ;
613 ;
614 ;
615 ;
616 ;
617 ;
618 ;
619 ;
620 ;
621 ;
622 ;
623 ;
624 ;
625 ;
626 ;
627 ;
628 ;
629 ;
630 ;
631 ;
632 ;
633 ;
634 ;
635 ;
636 ;
637 ;
638 ;
639 ;
640 ;
641 ;
642 ;
643 ;
644 ;
645 ;
646 ;
647 ;
648 ;
649 ;
650 ;
651 ;
652 ;
653 ;
654 ;
655 ;
656 ;
657 ;
658 ;
659 ;
660 ;
661 ;
662 ;
663 ;
664 ;
665 ;
666 ;
667 ;
668 ;
669 ;
670 ;
671 ;
672 ;
673 ;
674 ;
675 ;
676 ;
677 ;
678 ;
679 ;
680 ;
681 ;
682 ;
683 ;
684 ;
685 ;
686 ;
687 ;
688 ;
689 ;
690 ;
691 ;
692 ;
693 ;
694 ;
695 ;
696 ;
697 ;
698 ;
699 ;
700 ;
701 ;
702 ;
703 ;
704 ;
705 ;
706 ;
707 ;
708 ;
709 ;
710 ;
711 ;
712 ;
713 ;
714 ;
715 ;
716 ;
717 ;
718 ;
719 ;
720 ;
721 ;
722 ;
723 ;
724 ;
725 ;
726 ;
727 ;
728 ;
729 ;
730 ;
731 ;
732 ;
733 ;
734 ;
735 ;
736 ;
737 ;
738 ;
739 ;
740 ;
741 ;
742 ;
743 ;
744 ;
745 ;
746 ;
747 ;
748 ;
749 ;
750 ;
751 ;
752 ;
753 ;
754 ;
755 ;
756 ;
757 ;
758 ;
759 ;
760 ;
761 ;
762 ;
763 ;
7
```

Db	125	-----ARMYSKSDGTGSDDKA-----	142
Qy	421	DPLIQSPSPAVCEPPSPKHWSSVTSRTGSSGAKEMKLKGADGKTKIATPRGAAPPQOK	480
Db	143	-----KGADGKTKIATPRGAAPPQOK	163
Qy	481	GOANATRIPAKTPAPKTPPSSGEPKSGDRSGYSSPGSPGTPGSRSTPISLPTPTREP	540
Db	164	GOANATRIPAKTPAPKTPPSSGEPKSGDRSGYSSPGSPGTPGSRSTPISLPTPTREP	223
Qy	541	KKVAVRTPPKSPSSAKSRLQTAPVMPOLKNVKSIGSTENLKHPQGGKQVQIINKKLD	600
Db	224	KKVAVRTPPKSPSSAKSRLQTAPVMPOLKNVKSIGSTENLKHPQGGKQVQIINKKLD	283
Qy	601	LSNVQSKGSKDNIKHVPGGGSVQIIVYKVDLSKVTSKCGSLGNIHHKPGGGOVEVKSEK	660
Db	284	LSNVQSKGSKDNIKHVPGGGSVQIIVYKVDLSKVTSKCGSLGNIHHKPGGGOVEVKSEK	343
Qy	661	LDFKDRVQSKIGSLDNITTHVPGGNGKKEIETHKLTFFRENKAKTDHCAEIVYKSPVVSODT	720
Db	344	LDFKDRVQSKIGSLDNITTHVPGGNGKKEIETHKLTFFRENKAKTDHCAEIVYKSPVVSODT	403
Qy	721	SPRHLNSVSTGSDIMVDSPLATLADEVSAKAKOGL	758
Db	404	SPRHLNSVSTGSDIMVDSPLATLADEVSAKAKOGL	441

RESULT 9
 US-09-640-737-1
 Sequence 1, Application US/09640737
 GENERAL INFORMATION:
 APPLICANT: Mandelkow, Eva-Maria
 Mandelkow, Eckhard
 Lichtenberg-Kraag, Birgit
 Biernat, Jacek
 Drewes, Gerard
 Steiner, Barbara
 TITLE OF INVENTION: Novel Tools For The Diagnosis And
 Treatment Of Alzheimer's Disease
 NUMBER OF SEQUENCES: 1
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
 STREET: 233 South Waker Drive, 6300 Sears Tower
 CITY: Chicago
 STATE: Illinois
 COUNTRY: United States of America
 ZIP: 60606-6402
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Tape
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/640,737
 FILING DATE: 17-Aug-2000
 CLASSIFICATION: <unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: Joseph A. Williams, Jr.
 REGISTRATION NUMBER: 38,659
 REFERENCE/DOCKET NUMBER: 28384/32777
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312-474-6300
 TELEFAX: 312-484-0448
 TELEX: 25-3856
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 441 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-640-737-1

```
Query Match      53.4%; Score 2116.5; DB 20; Length 441;
Best Local Similarity 58.2%; Pred. No. 3.2e-108;
Matches 441; Conservative 0; Mismatches 0; Indels 317; Gaps 2;

QY 1 MAEPQEFVWEDHAGTYGLGDRKQGGYTHHQDEGDTDAGLKESPLQPTTDEGSEEPG 60
D 1 MAEPQEFVWEDHAGTYGLGDRKQGGYTHHQDEGDTDAGLKESPLQPTTDEGSEEPG 60
QY 61 SETSDAKSTPTAEDVTAPLVDGAPGKQAAQPHTEIPEGTTAEAGIGDTPSLEDAAG 120
D 61 SETSDAKSTPTAEDVTAPLVDGAPGKQAAQPHTEIPEGTTAEAGIGDTPSLEDAAG 120
QY 121 HVTQEPESGVVQEGFLREPGLSHQMSGMPGAPLLPEGPREATRQPSGTGPDTEG 180
D 121 HVTQ-----124
QY 181 GRHAPELLKHOLLGDLHQEGPPLKAGGKERPGSKKEEVEDRDVDDESSPDSPSKASPA 240
D 125 -----124
QY 241 QDGRPPOTAAREATSIIPGPAEGAIPLVDPLSKVSTEIPASEPDGPGSVGRAGQDAPLE 300
D 125 -----124
QY 301 FTFHVEITPNVQKEAHSEHGLGRAAFPGEPEARGPSLGEDTKADLPPESEKQPA 360
D 125 -----124
QY 361 AAPRGKPVSRVPLKARMVSKSDGTGSDDKKAKTSTRSSAKTLKNRCLSPKLPTPCSS 420
D 125 -----142
QY 421 DPLIQSSPAVCEPEPPSPKHVSSVTSRTGSSGAKEMKLGADGKTKIATPRGAAPGQK 480
D 143 -----163
QY 481 GOANATRIPAKTPAPKTPSSGPPPKSGDRSGYSSPGSGTTPGSRSPSLPTPTREP 540
D 164 GOANATRIPAKTPAPKTPSSGPPPKSGDRSGYSSPGSGTTPGSRSPSLPTPTREP 223
QY 541 KKVAVRTPPKSPSSAKSLQTAAPVMPDLKNVKSIGSTENLKHQPGGKQVQIINKKLD 600
D 224 KKVAVRTPPKSPSSAKSLQTAAPVMPDLKNVKSIGSTENLKHQPGGKQVQIINKKLD 283
QY 601 LSNVQSKGSKDNKIHVPGGSVQIYKPVDSLKVTSCGSLGNTHHKPGGQGVKSK 660
D 284 LSNVQSKGSKDNKIHVPGGSVQIYKPVDSLKVTSCGSLGNTHHKPGGQGVKSK 343
QY 661 LDFKDRVQSKIGSLDNITHVPGGNNKIIETHKLTFRENAKAKTDHGAIEVYKSPVVSQGT 720
D 344 LDFKDRVQSKIGSLDNITHVPGGNNKIIETHKLTFRENAKAKTDHGAIEVYKSPVVSQGT 403
QY 721 SPRHLSNVSTGSDIMVDSPLATLADEVASLAKOGL 758
D 404 SPRHLSNVSTGSDIMVDSPLATLADEVASLAKOGL 441
```

RESULT 10

```
US-09-640-738-1
: Sequence 1, Application US/09640738
: GENERAL INFORMATION:
: APPLICANT: Mandelkow, Eva-Maria
: Mandelkow, Eckhard
: Lichtenberg-Kraag, Birgit
: Biernat, Jacek
: Drewes, Gerard
: Steiner, Barbara
: TITLE OF INVENTION: Novel Tools For The Diagnosis And
: Treatment Of Alzheimer's Disease
: NUMBER OF SEQUENCES: 1
: CORRESPONDENCE ADDRESS:
```

```
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Tape
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/640,738
FILING DATE: 17-Aug-2000
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Joseph A. Williams, Jr.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 28384/32777
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-484-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 441 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-640-738-1
```

```
Query Match      53.4%; Score 2116.5; DB 20; Length 441;
Best Local Similarity 58.2%; Pred. No. 3.2e-108;
Matches 441; Conservative 0; Mismatches 0; Indels 317; Gaps 2;

QY 1 MAEPQEFVWEDHAGTYGLGDRKQGGYTHHQDEGDTDAGLKESPLQPTTDEGSEEPG 60
D 1 MAEPQEFVWEDHAGTYGLGDRKQGGYTHHQDEGDTDAGLKESPLQPTTDEGSEEPG 60
QY 61 SETSDAKSTPTAEDVTAPLVDGAPGKQAAQPHTEIPEGTTAEAGIGDTPSLEDAAG 120
D 61 SETSDAKSTPTAEDVTAPLVDGAPGKQAAQPHTEIPEGTTAEAGIGDTPSLEDAAG 120
QY 121 HVTQEPESGVVQEGFLREPGLSHQMSGMPGAPLLPEGPREATRQPSGTGPDTEG 180
D 121 HVTQ-----124
QY 181 GRHAPELLKHOLLGDLHQEGPPLKAGGKERPGSKKEEVEDRDVDDESSPDSPSKASPA 240
D 125 -----124
QY 241 QDGRPPOTAAREATSIIPGPAEGAIPLVDPLSKVSTEIPASEPDGPGSVGRAGQDAPLE 300
D 125 -----124
QY 301 FTFHVEITPNVQKEAHSEHGLGRAAFPGEPEARGPSLGEDTKADLPPESEKQPA 360
D 125 -----124
QY 361 AAPRGKPVSRVPLKARMVSKSDGTGSDDKKAKTSTRSSAKTLKNRCLSPKLPTPCSS 420
D 125 -----142
QY 421 DPLIQSSPAVCEPEPPSPKHVSSVTSRTGSSGAKEMKLGADGKTKIATPRGAAPGQK 480
D 143 -----163
QY 481 GOANATRIPAKTPAPKTPSSGPPPKSGDRSGYSSPGSGTTPGSRSPSLPTPTREP 540
D 164 GOANATRIPAKTPAPKTPSSGPPPKSGDRSGYSSPGSGTTPGSRSPSLPTPTREP 223
```

QY 541 KKVAVRTTPKSPSSAKSLQTAAPVMPDLKNVKSIGSTENLKHQPGGKQVQIINKKLD 600
|||||
Db 224 KKVAVRTTPKSPSSAKSLQTAAPVMPDLKNVKSIGSTENLKHQPGGKQVQIINKKLD 283
QY 601 LSNVQSKGSKDNKIKHVPVGGSGVQIYKPVDSLKSVTSKSGSLGNIHHKPGGQGVVEVKSEK 560
|||||
Db 284 LSNVQSKGSKDNKIKHVPVGGSGVQIYKPVDSLKSVTSKSGSLGNIHHKPGGQGVVEVKSEK 343
QY 661 LDFKDRVQSKIGSLDNTHTHVPVGGGKKIETHKLTFRENAKAKTDHGAETVYKSPVVSQGT 720
|||||
Db 344 LDFKDRVQSKIGSLDNTHTHVPVGGGKKIETHKLTFRENAKAKTDHGAETVYKSPVVSQGT 403
QY 721 SPRHLSNVSTGSDIMVDSPLATLADEVSASLAKOGL 758
|||||
Db 404 SPRHLSNVSTGSDIMVDSPLATLADEVSASLAKOGL 441
RESULT 11
US-10-066-810-1
; Sequence 1, Application US/10066810
; GENERAL INFORMATION:
; APPLICANT: Mandelkow, Eva-Maria
; Mandelkow, Eckhard
; Lichtenberg-Kraag, Birgit
; Biernat, Jacek
; Drewes, Gerard
; Steiner, Barbara
; TITLE OF INVENTION: Novel Tools For The Diagnosis And
; Treatment Of Alzheimer's Disease
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Waker Drive, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Tape
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/066,810
; FILING DATE: 04-Feb-2002
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Joseph A. Williams, Jr.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 28384/32777
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-484-0448
; TELE: 25-3856
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; TYPE: amino acids
; LENGTH: 441 amino acids
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-066-810-1

Query Match 53.4%; Score 2116.5; DB 24; Length 441;
Best Local Similarity 58.2%; Pred. No. 3.2e-108;
Matches 441; Conservative 0; Mismatches 0; Indels 317; Gaps 2;
QY 1 MAEPQEFVEMDHAGTYGLGDRKQGGYTMHQDQEGDTDAGLKESPLQPTPTDGSSEPG 60
|||||
Db 1 MAEPQEFVEMDHAGTYGLGDRKQGGYTMHQDQEGDTDAGLKESPLQPTPTDGSSEPG 60
|||||
QY 61 SETSDAKSTPTAEDVTAPLVEGAPGKQAAAPHTPEGTABEAGIGDTPSLEDAAG 120

Db 61 SETSDAKSTPTAEDVTAPLVEGAPGKQAAAPHTPEGTABEAGIGDTPSLEDAAG 120
|||||
QY 121 HVTQEPESGKVVQEGFLREPGLSHQMLMSGMPGAPLLPEGPREATRQPSGTGPTDTEG 180
|||||
Db 121 HVTQ----- 124
QY 181 GRHAPPELLHOLLGDLHOEGPPLKGAGGKERPGSKKEEVEDRDVDESSPDSPSPSKASPA 240
----- 124
Db 125 ----- 124
QY 241 QDGRPPQTAAREATSIPGFPAGGAIPLVDFLSKVSTETIPASEPDGVSVGRAKQDAPLE 300
----- 124
Db 125 ----- 124
QY 301 FTFHVEITPNVQKEQAHSEHLGRAAPGAPGEGPEARGLGDTKEADLPPESEKQPA 360
----- 124
Db 125 ----- 124
QY 361 AAPRGKPVSRVQLKARMVSKSDGTGDDKAKATSTRSSAKTLKNRPCLSPKLPPTPSS 420
|||||
Db 125 -----ARMVSKSDGTGDDKKA----- 142
QY 421 DPLIQSPSPAVCEPPSPKVVSSVTSRTGSSGAKEMKLGADCKTKIATPRGAAPGQK 480
|||||
Db 143 -----KGADCKTKIATPRGAAPGQK 163
QY 481 GOANATRIPAKTPAPKTPPSSGEPKSGDRSGYSSPGSPGTSRSTPSLPTPTPREP 540
|||||
Db 164 GOANATRIPAKTPAPKTPPSSGEPKSGDRSGYSSPGSPGTSRSTPSLPTPTPREP 223
QY 541 KKVAVRTTPKSPSSAKSLQTAAPVMPDLKNVKSIGSTENLKHQPGGKQVQIINKKLD 600
|||||
Db 224 KKVAVRTTPKSPSSAKSLQTAAPVMPDLKNVKSIGSTENLKHQPGGKQVQIINKKLD 283
QY 601 LSNVQSKGSKDNKIKHVPVGGSGVQIYKPVDSLKSVTSKSGSLGNIHHKPGGQGVVEVKSEK 660
|||||
Db 284 LSNVQSKGSKDNKIKHVPVGGSGVQIYKPVDSLKSVTSKSGSLGNIHHKPGGQGVVEVKSEK 343
QY 661 LDFKDRVQSKIGSLDNTHTHVPVGGGKKIETHKLTFRENAKAKTDHGAETVYKSPVVSQGT 720
|||||
Db 344 LDFKDRVQSKIGSLDNTHTHVPVGGGKKIETHKLTFRENAKAKTDHGAETVYKSPVVSQGT 403
QY 721 SPRHLSNVSTGSDIMVDSPLATLADEVSASLAKOGL 758
|||||
Db 404 SPRHLSNVSTGSDIMVDSPLATLADEVSASLAKOGL 441
RESULT 12
US-10-076-708-2
; Sequence 2, Application US/10076708
; GENERAL INFORMATION:
; APPLICANT: Sharma, Satish
; APPLICANT: Rank, Kenneth
; TITLE OF INVENTION: Assays for Accessing Aa-Tau Aggregation
; FILE REFERENCE: 6322
; CURRENT APPLICATION NUMBER: US/10/076,708
; CURRENT FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 2
; LENGTH: 441
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-076-708-2

Query Match 53.4%; Score 2116.5; DB 24; Length 441;
Best Local Similarity 58.2%; Pred. No. 3.2e-108;
Matches 441; Conservative 0; Mismatches 0; Indels 317; Gaps 2;
QY 1 MAEPQEFVEMDHAGTYGLGDRKQGGYTMHQDQEGDTDAGLKESPLQPTPTDGSSEPG 60
|||||
Db 1 MAEPQEFVEMDHAGTYGLGDRKQGGYTMHQDQEGDTDAGLKESPLQPTPTDGSSEPG 60
|||||

QY 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAQAQPHTEIPEGTTAEAGIGDTPSLEDAAG 120
|||||
Db 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAQAQPHTEIPEGTTAEAGIGDTPSLEDAAG 120
|||||
QY 121 HVTQEPESGKVVQEGFLREPGLSHQLMGMPGAPLLPEGPREATROPSTGTPEDTEG 180
|||||
Db 121 HVTQ----- 124
QY 181 GRHAPELLKHLLGLDGHQEGPLKAGGKERPGSKEEVEDRDVDDESSPDSPSPSKASPA 240
Db 125 ----- 124
QY 241 QDGRPQTAAAREATSIIPGPAEGAIPLPVDFLSKVSTEIPASEPDGPGSVGRAKGDQAPLE 300
Db 125 ----- 124
QY 301 FTFHVEITPNVQKEQAHEEHLGRAAPGAPGEGPEARGPSLGEDTKEADLPPESEKQPA 360
Db 125 ----- 124
QY 361 AAPRGKPVSRVPQLKARVSKSDGTGDDKAKTSTRSSAKTLKRNCLSPKLPPTPGSS 420
Db 125 -----ARMVSKSDGTGDDKKA----- 142
QY 421 DPLIQSSPAVCEPPSPKHVSSVTSRTGSSGAKEMKLGADGKTKIATPRGAAPPQK 480
Db 143 -----KGADGKTKIATPRGAAPPQK 163
QY 481 GOANATRIPAKTPAPKTPSSGEPKSGDRSGYSSPGSPGTPGSRSTPSLPTPTPREP 540
Db 164 GOANATRIPAKTPAPKTPSSGEPKSGDRSGYSSPGSPGTPGSRSTPSLPTPTPREP 223
QY 541 KKVAVRTTPKSPSSAKSRLQTAPVMPDLKNVKSIGSTENLKHQPGGKQVQIINKKLD 600
Db 224 KKVAVRTTPKSPSSAKSRLQTAPVMPDLKNVKSIGSTENLKHQPGGKQVQIINKKLD 283
QY 601 LSNVQSKGSKDNKIHVPVGGGVSQIYKPVDSLKVTSKCSGLGNIHHKPGGGQVEVKSEK 660
Db 284 LSNVQSKGSKDNKIHVPVGGGVSQIYKPVDSLKVTSKCSGLGNIHHKPGGGQVEVKSEK 343
QY 661 LDFKDRVQSKIGSLDNITHVPGGNGKIIETHKLTFRENAKAKTDHGAEIYKSPVVSQGT 720
Db 344 LDFKDRVQSKIGSLDNITHVPGGNGKIIETHKLTFRENAKAKTDHGAEIYKSPVVSQGT 403
QY 721 SPRHLSNVSTGSDMVDSPQLATLADEVSAKLAKOGL 758
Db 404 SPRHLSNVSTGSDMVDSPQLATLADEVSAKLAKOGL 441

RESULT 13

US-60-271-102-2
; Sequence 2, Application US/60271102
; GENERAL INFORMATION:
; APPLICANT: Sharma, Satish
; APPLICANT: Rank, Kenneth
; TITLE OF INVENTION: Assays for Accessing AA-Tau Aggregation
; FILE REFERENCE: 6322
; CURRENT APPLICATION NUMBER: US/60/271,102
; CURRENT FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 441
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-271-102-2
Query Match 53.4%; Score 2116.5; DB 27; Length 441;
Best Local Similarity 58.2%; Pred. No. 3.2e-108;
Matches 441; Conservative 0; Mismatches 0; Indels 317; Gaps 2;
QY 1 MAEPQREFEVEDHAGTYGLGDRKQGGYTMHQDQGGDTAGLKESPLQPTPTEDGSEEPG 60

RESULT 14

US-09-142-613-1
; Sequence 1, Application US/09142613A
; GENERAL INFORMATION:
; APPLICANT: ISHIGURO, Koichi
; APPLICANT: SATO, Kazuki
; APPLICANT: PARK, Jun-Mi
; APPLICANT: UCHIDA, Tsuneko
; APPLICANT: IMAHORI, Kazutomo
; TITLE OF INVENTION: ANTI-PHOSPHORYLATED TAU PROTEIN ANTIBODIES AND METHODS
; TITLE OF INVENTION: FOR DETECTING ALZHEIMER'S DISEASE WITH THE USE OF THE
; FILE REFERENCE: 98-0997*/LC(WMC)/1416
; CURRENT APPLICATION NUMBER: US/09/142,613A
; CURRENT FILING DATE: 1999-04-19
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 441
; TYPE: PRT
; ORGANISM: Homo sapiens

Db 344 LDFKDRVQSKIGSLDNITHVPGGNKKIETHKLTVRENAKAKTDHGAELVYKSPVVGDT 403
QY 721 SPRHLSNVSTGSDWVDSFQATLADSVASLAKQGL 758
Db 404 SPRHLSNVSTGSDWVDSFQATLADSVASLAKQGL 441

Search completed: January 28, 2003, 14:37:25
Job time : 202.488 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:16:13 : Search time 22.1541 seconds
(without alignments)
2701.343 Million cell updates/sec

Title: US-09-904-987-5

Perfect score: 3965

Sequence: 1 MAEPQFEVMEHDHAGTYGL.....SPQLATLADEVSLAKQGL 758

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 394613 seqs, 78952108 residues

Total number of hits satisfying chosen parameters: 394613

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1590	40.1	391	6	US-10-308-150-23
2	1400.5	35.3	340	5	US-09-724-676-49498
3	1400.5	35.3	340	5	US-09-724-676A-49498
4	1214	30.6	309	5	US-09-724-676-49497
5	1214	30.6	309	5	US-09-724-676A-49497
6	951.5	24.0	1856	5	US-09-724-676-51334
7	951.5	24.0	1856	5	US-09-724-676A-51334
8	891	22.5	502	5	US-09-724-676-51337
9	891	22.5	502	5	US-09-724-676A-51337
10	811	20.5	1825	5	US-09-724-676-51331
11	811	20.5	1825	5	US-09-724-676A-51331
12	750.5	18.9	471	5	US-09-724-676-51332
13	750.5	18.9	471	5	US-09-724-676A-51332
14	657	16.6	125	6	US-10-070-611-1
15	621.5	15.7	1125	6	US-10-100-957A-152
16	621.5	15.7	1610	6	US-10-100-957A-22
17	593	15.0	112	6	US-10-247-853-1
18	482.5	12.2	1011	1	PCT-US02-32851-39
19	470.5	11.9	94	6	US-10-070-611-10
20	462.5	11.7	94	6	US-10-070-611-13
21	456	11.5	91	6	US-10-070-611-12
22	455.5	11.5	94	6	US-10-070-611-11
23	415	10.5	1068	1	PCT-US02-32851-44
24	364	9.2	67	6	US-10-308-150-1
25	362.5	9.1	1693	5	US-09-724-676-51333
26	362.5	9.1	1693	5	US-09-724-676A-51333

ALIGNMENTS

RESULT 1

US-10-308-150-23

: Sequence 23, Application US/10308150

: GENERAL INFORMATION:

: APPLICANT: VANDEMEEREN, MARC; MERCKEN, MARC;

: VANMECHELEN, EUGEN;

: VAN DE VOORDE, ANDRE

: TITLE OF INVENTION: MONOCLONAL ANTIBODIES

: DIRECTED AGAINST THE MICROTUBULE-ASSOCIATED
: PROTEIN TAU, HYBRIDOMAS SECRETING THESE
: ANTIBODIES, ANTIGEN RECOGNITION BY THESE

: NUMBER OF SEQUENCES: 24

: CORRESPONDENCE ADDRESSES:

: ADDRESSEE: BIERMAN & MUSERLIAN

: STREET: 600 THIRD AVENUE

: CITY: NEW YORK

: STATE: NEW YORK

: COUNTRY: USA

: ZIP: 10016

: COMPUTER READABLE FORM:

: MEDIUM TYPE: FLOPPY DISK

: COMPUTER: IBM PC COMPATIBLE

: OPERATING SYSTEM: PC-DOS/MS-DOS

: SOFTWARE: ASCII

: CURRENT APPLICATION DATA:

: APPLICATION NUMBER: US/10308,150

: FILING DATE: 02-Nov-2002

: PRIOR APPLICATION DATA:

: APPLICATION NUMBER: <Unknown>

: FILING DATE: 07-JUL-1999

: APPLICATION NUMBER: 09/348,952

: FILING DATE: 19-JAN-1995

: APPLICATION NUMBER: 08/403,917

: FILING DATE: 19-JAN-1995

: APPLICATION NUMBER: 08/403,916

: FILING DATE: 13-JUN-1994

: APPLICATION NUMBER: 08/244,951

: FILING DATE: 10-DEC-1993

: APPLICATION NUMBER: PCT/EP93/03499

: FILING DATE: 14-DEC-1992

: APPLICATION NUMBER: EP/92/403403.6

: FILING DATE: 14-DEC-1992

: ATTORNEY/AGENT INFORMATION:

: NAME: CHARLES A. MUSERLIAN

: REGISTRATION NUMBER: 19,683

: REFERENCE/DOCKET NUMBER: 410.003-1

: TELECOMMUNICATION INFORMATION:

: TELEPHONE: (212) 661-8000

: TELEFAX: (212) 661-8002

Sequence 51336, A
Sequence 51336, A
Sequence 1, Appli
Sequence 72859, A
Sequence 72859, A
Sequence 7237, Ap
Sequence 202, App
Sequence 202, App
Sequence 82068, A
Sequence 82068, A
Sequence 82069, A
Sequence 82069, A
Sequence 72, Appl
Sequence 405, App
Sequence 492, App
Sequence 12693, A
Sequence 75580, A
Sequence 75580, A
Sequence 139, App

27 302 7.6 339 5 US-09-724-676-51336
28 302 7.6 339 5 US-09-724-676A-51336
29 293 7.4 1374 1 PCT-US02-34846-1
30 288 7.3 1610 5 US-09-724-676-72859
31 288 7.3 1610 5 US-09-724-676A-72859
32 287.5 7.3 99 5 US-09-513-999C-7237
33 285.5 7.2 1806 1 PCT-US02-29560-202
34 285.5 7.2 1806 6 US-10-245-882-202
35 285 7.2 1054 5 US-09-724-676-82068
36 285 7.2 1054 5 US-09-724-676A-82068
37 285 7.2 1149 5 US-09-724-676-82069
38 285 7.2 1149 5 US-09-724-676A-82069
39 283.5 7.2 1466 6 US-10-257-021-72
40 283.5 7.2 1745 5 US-09-949-002-405
41 283.5 7.2 1771 5 US-09-949-002-492
42 282 7.1 1325 6 US-10-203-138A-12693
43 278.5 7.0 1636 5 US-09-724-676-75580
44 278.5 7.0 1636 5 US-09-724-676A-75580
45 277 7.0 1604 7 US-60-423-552-139

; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 391
; TYPE: Amino Acid
; STRANDEDNESS: Unknown
; TOPOLOGY: Unknown
; SEQUENCE DESCRIPTION: SEQ ID NO: 23:
US-10-308-150-23

Query Match 40.1%; Score 1590; DB 6; Length 391;
Best Local Similarity 46.4%; Pred. No. 1.7e-64;
Matches 352; Conservative 0; Mismatches 0; Indels 406; Gaps 4;
QY 1 MAEPRQEFVEMDHACTYGLGRKDGQYTMHQDQGDTHDAGLKESPLQTPTEDESGEPPG 60
Db 40 MAEPRQEFVEMDHACTYGLGRKDGQYTMHQDQGDTHDAGLKESPLQTPTEDESGEPPG 83
QY 61 SETSDAKSTPTAEDVTAPLVDEGAPGKAAQPHTEIPGTTAERAGIGDTPSLDEDAAG 120
Db 84 -----AEEAGIGDTPSLDEDAAG 101
QY 121 HVTQEPESGKVVQEGFLREPPGGLSHQLMSGMPGAPLLPEGPREATRQPSGTGPDTEG 180
Db 102 HVTQ----- 105
QY 181 GRHAPELLKHQLLDLHQEGPPLKGAGKERPGSKKEEVEDRDVDESSPDSPSPKASPA 240
Db 106 ----- 105
QY 241 QDGRPPQTAAREATSTPGFPAEGAIFLPVDFLSKVSTETPASEPDGPGSVGRAKQGDAPLE 300
Db 106 ----- 105
QY 301 FTFHVEITPVQKEQAHESEHLGRAAPGAPGEGPEARGPSLGEDTKADLPPEPSEKQPA 360
Db 106 ----- 105
QY 361 AAPRGKPVSRVPOLKARWVSKSDGTGSDDKAKKTSTRSSAKTLKNRPCLSPKLPPTPGSS 420
Db 106 -----ARMVSKSDGTGSDDKKA----- 123
QY 421 DPLIQSSPAVCPEPPSSPKHVSVTSRTGSSGAKEMKLKADGKTKIATPRGAAPPQOK 480
Db 124 -----KGADGKTKIATPRGAAPPQOK 144
QY 481 QGANATRIAPKTPAPKTPSSGEPKSGDRSGYSSPGSPGTPGSRSTPSPSLPTPTREP 540
Db 145 QGANATRIAPKTPAPKTPSSGEPKSGDRSGYSSPGSPGTPGSRSTPSPSLPTPTREP 204
QY 541 KKVAVRTPPKSPSSAKSRLOTAPVPMPLKNVSKSIGSTENLKHQPGGKVOIINKLKD 600
Db 205 KKVAVRTPPKSPSSAKSRLOTAPVPMPLKNVSKSIGSTENLKHQPGGK----- 255
QY 601 LSNVQSCGSKDNKIHVPGGGSQVIVYKPVDSLKVTSKCSLGNIIHKKPGGQVEVKSEK 660
Db 256 -----VQIVYKPVDSLKVTSKCSLGNIIHKKPGGQVEVKSEK 293
QY 661 LDFKDRVQSKIGSLDNITHVPGGNGKKIETHKLTFFRENAAKTDHGAIEVYKSPVVSGET 720
Db 294 LDFKDRVQSKIGSLDNITHVPGGNGKKIETHKLTFFRENAAKTDHGAIEVYKSPVVSGET 353
QY 721 SPRHLSNVSTGSDIMVDSPLATLADSVASLAKOGL 758
Db 354 SPRHLSNVSTGSDIMVDSPLATLADSVASLAKOGL 391

RESULT 2
US-09-724-676-49498
; Sequence 49498, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Comugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Comugen

; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 49498
; LENGTH: 340
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-49498
Query Match 35.3%; Score 1400.5; DB 5; Length 340;
Best Local Similarity 75.1%; Pred. No. 4.8e-56;
Matches 280; Conservative 5; Mismatches 7; Indels 81; Gaps 3;
QY 353 EPSEKQPAAPGRKPVSRVPOL-----KARMVSKSDGTGSDDKAKKTSTRSSAKT 403
Db 19 QPAEAEAG-----IGDTPSLEDEAAGHVTAQRMVSKSDGTGSDDKKA----- 62
QY 404 LKNRPCLSPKLPPTGSSDPLIQSSPAVCPEPPSPKHVSVTSRTGSSGAKEMKLKAD 463
Db 63 -----KGAD 66
QY 464 GKTKIATPRGAAPGKQGANATRIAPKTPAPKTPSSGEPKSGDRSGYSSPGSPCTP 523
Db 67 GKTKIATPRGAAPGKQGANATRIAPKTPAPKTPSSGEPKSGDRSGYSSPGSPCTP 126
QY 524 GSRSTPSPSLPTPTREPKKVAVRTPPKSPSSAKSRLOTAPVPMPLKNVSKSIGSTENL 583
Db 127 GSRSTPSPSLPTPTREPKKVAVRTPPKSPSSAKSRLOTAPVPMPLKNVSKSIGSTENL 186
QY 584 KHQPGGKVOIINKLDSLNVQSKCGSKDNKIHVPGGGSQVIVYKPVDSLKVTSKCSL 643
Db 187 KHQPGGKVOIINKLDSLNVQSKCGSKDNKIHVPGGGSQVIVYKPVDSLKVTSKCSL 246
QY 644 NIHHKPGGQVEVKSEKLDKDRVQSKIGSLDNITHVPGGNGKKIETHKLTFFRENAK 703
Db 247 NIHHKPGGQVEVKSEKLDKDRVQSKIGSLDNITHVPGGNGKKIETHKLTFFRENAK 306
QY 704 DHGAEIVYKSPV 716
Db 307 DHGAEIVYKSPV 319

RESULT 3
US-09-724-676A-49498
; Sequence 49498, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Comugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Comugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 49498
; LENGTH: 340
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-49498

Query Match 35.3%; Score 1400.5; DB 5; Length 340;
Best Local Similarity 75.1%; Pred. No. 4.8e-56;
Matches 280; Conservative 5; Mismatches 7; Indels 81; Gaps 3;
QY 353 EPSEKQPAAPGRKPVSRVPOL-----KARMVSKSDGTGSDDKAKKTSTRSSAKT 403
Db 19 QPAEAEAG-----IGDTPSLEDEAAGHVTAQRMVSKSDGTGSDDKKA----- 62
QY 404 LKNRPCLSPKLPPTGSSDPLIQSSPAVCPEPPSPKHVSVTSRTGSSGAKEMKLKAD 463
Db 63 -----KGAD 66
QY 464 GKTKIATPRGAAPGKQGANATRIAPKTPAPKTPSSGEPKSGDRSGYSSPGSPCTP 523

Db 67 GKTATPRGAAPPQGGKQANATRIAPKTPPSSSEPPKSGDRSGYSSPGSGTP 126
QY 524 GSRSTPSLPTPTREPCKVAVVTRPPKSPSSAKSLRLQAPVMPDLKNVSKIGSTENL 583
Db 127 GSRSTPSLPTPTREPCKVAVVTRPPKSPSSAKSLRLQAPVMPDLKNVSKIGSTENL 186
QY 584 KHOPGGKQVQIINKLDLSNVQSKGSKDNIIHVPGGGQVQIYKPVLDLSKVTSCGSLG 643
Db 187 KHOPGGKQVQIINKLDLSNVQSKGSKDNIIHVPGGGQVQIYKPVLDLSKVTSCGSLG 246
QY 644 NIHKPGGGQVEVSKELDFKDRVQSKIGSLDNITHVPGGNNKKTETHKLTFRENAKAKT 703
Db 247 NIHKPGGGQVEVSKELDFKDRVQSKIGSLDNITHVPGGNNKKTETHKLTFRENAKAKT 306
QY 704 DHGAEIYKSPV 716
Db 307 DHGAEIYKSPV 319

RESULT 4

US-09-724-676-49497
; Sequence 49497, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 49497
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-49497

Query Match 30.6%; Score 1214; DB 5; Length 309;
Best Local Similarity 66.8%; Pred. No. 1e-47;
Matches 249; Conservative 5; Mismatches 7; Indels 112; Gaps 4;

QY 353 EPSEKQPAAPRGKPVSRVQPL-----KARVMSKSKDGTGSDDKKAKTSTRSSAKT 403
Db 19 QPAEAEAG-----IGDTPSLEDEAAGHVTAQARMVSKSKDGTGSDDKKA----- 62
QY 404 LKNRPCLSPKLPFGSSDPLIQSPSPAVCPPEPPSPKHVSVTSRTGSSGAKEMKLGAD 463
Db 63 -----KGAD 66
QY 464 GKTATPRGAAPPQGGKQANATRIAPKTPPSSSEPPKSGDRSGYSSPGSGTP 523
Db 67 GKTATPRGAAPPQGGKQANATRIAPKTPPSSSEPPKSGDRSGYSSPGSGTP 126
QY 524 GSRSTPSLPTPTREPCKVAVVTRPPKSPSSAKSLRLQAPVMPDLKNVSKIGSTENL 583
Db 127 GSRSTPSLPTPTREPCKVAVVTRPPKSPSSAKSLRLQAPVMPDLKNVSKIGSTENL 186
QY 584 KHOPGGKQVQIINKLDLSNVQSKGSKDNIIHVPGGGQVQIYKPVLDLSKVTSCGSLG 643
Db 187 KHOPGGK-----VQIYKPVLDLSKVTSCGSLG 215
QY 644 NIHKPGGGQVEVSKELDFKDRVQSKIGSLDNITHVPGGNNKKTETHKLTFRENAKAKT 703
Db 216 NIHKPGGGQVEVSKELDFKDRVQSKIGSLDNITHVPGGNNKKTETHKLTFRENAKAKT 275
QY 704 DHGAEIYKSPV 716
Db 276 DHGAEIYKSPV 288

RESULT 5

US-09-724-676A-49497
; Sequence 49497, Application US/09724676A

; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 49497
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-49497

Query Match 30.6%; Score 1214; DB 5; Length 309;
Best Local Similarity 66.8%; Pred. No. 1e-47;
Matches 249; Conservative 5; Mismatches 7; Indels 112; Gaps 4;

QY 353 EPSEKQPAAPRGKPVSRVQPL-----KARVMSKSKDGTGSDDKKAKTSTRSSAKT 403
Db 19 QPAEAEAG-----IGDTPSLEDEAAGHVTAQARMVSKSKDGTGSDDKKA----- 62
QY 404 LKNRPCLSPKLPFGSSDPLIQSPSPAVCPPEPPSPKHVSVTSRTGSSGAKEMKLGAD 463
Db 63 -----KGAD 66
QY 464 GKTATPRGAAPPQGGKQANATRIAPKTPPSSSEPPKSGDRSGYSSPGSGTP 523
Db 67 GKTATPRGAAPPQGGKQANATRIAPKTPPSSSEPPKSGDRSGYSSPGSGTP 126
QY 524 GSRSTPSLPTPTREPCKVAVVTRPPKSPSSAKSLRLQAPVMPDLKNVSKIGSTENL 583
Db 127 GSRSTPSLPTPTREPCKVAVVTRPPKSPSSAKSLRLQAPVMPDLKNVSKIGSTENL 186
QY 584 KHOPGGKQVQIINKLDLSNVQSKGSKDNIIHVPGGGQVQIYKPVLDLSKVTSCGSLG 643
Db 187 KHOPGGK-----VQIYKPVLDLSKVTSCGSLG 215
QY 644 NIHKPGGGQVEVSKELDFKDRVQSKIGSLDNITHVPGGNNKKTETHKLTFRENAKAKT 703
Db 216 NIHKPGGGQVEVSKELDFKDRVQSKIGSLDNITHVPGGNNKKTETHKLTFRENAKAKT 275
QY 704 DHGAEIYKSPV 716
Db 276 DHGAEIYKSPV 288

RESULT 6

US-09-724-676-51334
; Sequence 51334, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 51334
; LENGTH: 1856
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-51334

Query Match 24.0%; Score 951.5; DB 5; Length 1856;
Best Local Similarity 31.5%; Pred. No. 3.9e-35;
Matches 307; Conservative 125; Mismatches 281; Indels 261; Gaps 39;

QY 5 RQEFVEMDHAGTYGLGDRKQGGYTHHQDQEGDTHAGLKESPLQTPTEDESGSEPGSETS 64
Db 924 KDEFV-DKEASAHISGDK---SGLSKFEFDOEKKAN-----DRLDTVLEK-SEEHADSKE 973
QY 65 DAKSTPTAEDYTAPLVDEGAPGKQAAQPHTEIPGTTAEEA--GIGDTPSLEDEAAGHV 122

```
Db 974 HAKTEAGD-EIETFGVLYEQALAK-DLSIPTDASSEKAEKGLSSVPEI----- 1023
Qy 123 TQPESEKVVQEGF-----LREPGLSHQMSGM-----PGAPLLPEGPREA 166
Db 1024 -AEVPSKVEOGLDFAVQGLDVKISDFG-----QMASGLNIDRRATELKEATQDM 1076
Qy 167 TROPSTGTPEDT-----EGGRHAPELLKHQLGQD--LHQEGPPLK---GAGGK 209
Db 1077 T--PSSKAPQADAFMGVESHMKEGTKVSETEVKQKAPDLVHQEAVDKAESYESSGE 1134
Qy 210 ERPGSKEEVEDRDVDSESPQDS-----PP--SKASPAQDGRP----- 245
Db 1135 HESLTMESLKADGKKTSPSSLIQDEIAVKLSVEIPCPVAVSEADLATDERADVQMEF 1194
Qy 246 ---PQTAAREATSIPIGPAEAGIPLVPDLFSLKSVTEIPASEPDGPGSVGRAGQDAPLEFT 302
Db 1195 IQGPKESEKTPDISITPSDVAEPLHETIVSE-PAEQSEEEE-----IEAQGEYDKLLFR 1249
Qy 303 F-----HVEITPNVOK-----EQAHSEHGLGRAAPFG 329
Db 1250 SDTLQITDLGVSAREFVETCPSEHKGVIESTIEDDTITVVQTTTDEGESGSHSVRF 1309
Qy 330 APGEGPEARGPSLGEDTKADLPPESEKQ-----PAAAPRGKPVSRVPOLKARVMS 380
Db 1429 EKTGRGRISTPERKVAKEPSTVSRDEVRKKAIVYKAEAKKTEVQAHSRSPKFIKPA 1488
Qy 437 ---SSPKHVSVTSRTGSSGA-----KEMKLGADGKTKIATPRGAAP-----PG 478
Db 1489 IKYTRPHLSCVKKRKTAAAGGESALAPSVFKQAKDKVSDGVTSPKRSRSLPPSSILPP 1548
Qy 479 OKG-----QANATRIPAKTP-----PAPKTPPSS---GEPPKSG 509
Db 1549 RRGVSGDRDENSEFSLNSSLSSARRTTRSEPIRRAGKSGTSTPTTGTATPTGTPPSS 1608
Qy 510 DRSGYSSPGSGPTGSRSTPSLTPP-----TREPKKVAVVTRPPKSPSSAKSRLOATP 564
Db 1609 SR---TPGTGTP-SYRTPHTPGTKSAILVPSEKKVAIIRTPPKSPATPK-QLRLIN 1662
Qy 565 VPMPLKNVSKIGSTENLKHQGGKQVQIINKKLDLSNVQSCGSKDNKIHVPVGGG 624
Db 1663 QPLPDLKNVSKIGSTDNKIKYQPKGGQVRILNKKIDFSKVQSCGSKDNKIHAGGNNVQ 1722
Qy 625 IVYKPVDLKSVTSKCSLGNIIHHPGGGQVEVSEKLDKDRVQSKIGSLDNITHVPGGG 684
Db 1723 IVTKKIDLSHVTSCGSLNIRHRRGGGRVKIESVKLDFEKAQAQKVGSLDNAHVPGGG 1782
Qy 685 NKKIETHKLTIFRENAKAKTDHGAIEIVYKSPVVSGBTSPRHLNSVSTGSDIMVDSPOLAT 744
Db 1783 NVKIDSQKLNFRHAKARVDHGAIEITQPGRRSVASPRRLSNVSSGSGINLLESPQAT 1842
Qy 745 LADEVASLAKOGL 758
Db 1843 LAEDVTAALAKOGL 1856
```

RESULT 7

```
US-09-724-676A-51334
; GENERAL INFORMATION:
; APPLICANT: Comugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Comugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
```

```
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 51334
; LENGTH: 1856
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-51334
```

Query Match 24.0%; Score 951.5; DB 5; Length 1856;

Best Local Similarity 31.5%; Pred. No. 3.9e-35;

Matches 307; Conservative 125; Mismatches 281; Indels 261; Gaps 39;

```
Qy 5 ROBEVEVEDIAGTYGLGRKQGGYTHMQOQEGDTAGLKESPLQDTEDGSESEPSSETS 64
Db 924 KDEFSV-DKEASAIHSGD---SGLSEKFEQDEKAN-----DRLDTVLEK-SEHADSK 973
Qy 65 DAKSTPTAEDVTAPLVDEGAPGKQAAAPHTETPEGTAAEA--GIGDTPSLEDEAGHV 122
Db 974 HAKTEAGD-EIETFGVLYEQALAK-DLSIPTDASSEKAEKGLSSVPEI----- 1023
Qy 123 TQPESEKVVQEGF-----LREPGLSHQMSGM-----PGAPLLPEGPREA 166
Db 1024 -AEVPSKVEOGLDFAVQGLDVKISDFG-----QMASGLNIDRRATELKEATQDM 1076
Qy 167 TROPSTGTPEDT-----EGGRHAPELLKHQLGQD--LHQEGPPLK---GAGGK 209
Db 1077 T--PSSKAPQADAFMGVESHMKEGTKVSETEVKQKAPDLVHQEAVDKAESYESSGE 1134
Qy 210 ERPGSKEEVEDRDVDSESPQDS-----PP--SKASPAQDGRP----- 245
Db 1135 HESLTMESLKADGKKTSPSSLIQDEIAVKLSVEIPCPVAVSEADLATDERADVQMEF 1194
Qy 246 ---PQTAAREATSIPIGPAEAGIPLVPDLFSLKSVTEIPASEPDGPGSVGRAGQDAPLEFT 302
Db 1195 IQGPKESEKTPDISITPSDVAEPLHETIVSE-PAEQSEEEE-----IEAQGEYDKLLFR 1249
Qy 303 F-----HVEITPNVOK-----EQAHSEHGLGRAAPFG 329
Db 1250 SDTLQITDLGVSAREFVETCPSEHKGVIESTIEDDTITVVQTTTDEGESGSHSVRF 1309
Qy 330 APGEGPEARGPSLGEDTKADLPPESEKQ-----PAAAPRGKPVSRVPOLKARVMS 380
Db 1310 AALQOPEVERRPSHDEEVEEAEAAQAEKPDGSPAPASPEREEVA-LSEKTYETD 1368
Qy 381 KSKDGTGSDQ-----LSPKLTGSSDPL-----KRAKTSRSL-ARTLKNRP 408
Db 1369 DYKDETTIDDSIMDADSLWDTQDDRSIMTEQLETTIPKEEKAKEARRSLEKHKRKP 1428
Qy 409 C-----LSPKLTGSSDPL-----IQPSSPA--VCEPP 436
Db 1429 FKTGRGRISTPERKVAKEPSTVSRDEVRKKAIVYKAEAKKTEVQAHSRSPKFIKPA 1488
Qy 437 ---SSPKHVSVTSRTGSSGA-----KEMKLGADGKTKIATPRGAAP-----PG 478
Db 1489 IKYTRPHLSCVKKRKTAAAGGESALAPSVFKQAKDKVSDGVTSPKRSRSLPPSSILPP 1548
Qy 479 OKG-----QANATRIPAKTP-----PAPKTPPSS---GEPPKSG 509
Db 1549 RRGVSGDRDENSEFSLNSSLSSARRTTRSEPIRRAGKSGTSTPTTGTATPTGTPPSS 1608
Qy 510 DRSGYSSPGSGPTGSRSTPSLTPP-----TREPKKVAVVTRPPKSPSSAKSRLOATP 564
Db 1609 SR---TPGTGTP-SYRTPHTPGTKSAILVPSEKKVAIIRTPPKSPATPK-QLRLIN 1662
Qy 565 VPMPLKNVSKIGSTENLKHQGGKQVQIINKKLDLSNVQSCGSKDNKIHVPVGGG 624
Db 1663 QPLPDLKNVSKIGSTDNKIKYQPKGGQVRILNKKIDFSKVQSCGSKDNKIHAGGNNVQ 1722
Qy 625 IVYKPVDLKSVTSKCSLGNIIHHPGGGQVEVSEKLDKDRVQSKIGSLDNITHVPGGG 684
Db 1723 IVTKKIDLSHVTSCGSLNIRHRRGGGRVKIESVKLDFEKAQAQKVGSLDNAHVPGGG 1782
Qy 685 NKKIETHKLTIFRENAKAKTDHGAIEIVYKSPVVSGBTSPRHLNSVSTGSDIMVDSPOLAT 744
Db 1783 NVKIDSQKLNFRHAKARVDHGAIEITQPGRRSVASPRRLSNVSSGSGINLLESPQAT 1842
Qy 745 LADEVASLAKOGL 758
Db 1843 LAEDVTAALAKOGL 1856
```


Db	924	KDQSV-DKEASAHISGDK---SGLSKEDFOQKKAN-----DRJDTVLEK-SEBHADSKS	973
Qy	65	DAKSTTAEDVTAPLVDEGAPGKQAAQPHTEIEPGTTAEBA--GIGDTPTSLEDEAAGHV	122
Db	974	HAKKTEAGD-EIETFTGLGVTVEQALAK-DLSIPTDASSEKAENGLSSVPEI-----	1023
Qy	123	TQEPESGKVUVQEGF-----LRPGPGLSHQLMSCM-----PGAPLLPEGPREA	166
Db	1024	-AEVEFSKVEQGLDFAVOGQLDVAKISDFG-----QMASGLDIDRRATELKLLEATQDM	1076
Qy	167	TROPSTGTGPEDT-----EGGRHAPELLKHOLLGD--LHQEGPPLK---GAGGK	209
Db	1077	T---PSSKAQOEADAPMGVESGHKKEGTKVSETEVKQYAKPDLVHQEAVDKESYESGGE	1134
Qy	210	ERPGSKEEVEDRDVDDESSPQDS-----PP--SKAPADQGRP-----	245
Db	1135	HESLTMESLKADEGKETSPESLSIQDEITAVKLSVEIPCPPAVSEADLATDERADVOMEF	1194
Qy	246	---POTAAAREANTSIPGFAEAGALPVDVFLSKVSTEIPIASPDGPGSVGRAKGQDAPLEET	302
Db	1195	IQGPKEESKETPDISITPSDVAEPDHEITVSE-PAEIQOSEEEE---IEAQGEYDKLULFR	1249
Qy	303	F-----HVEITPNVQK-----EQARSEEHGLGRAAFPG	329
Db	1250	SDTLQITDLGVSGAREFVETCPSEHKGVIESVWTIEDFTTVVQTTTDEGESSHVSFR	1309
Qy	330	APGEGPEARGPSLGEDTKEADLPERSEKQ-----PAAAPRGKPVSRVPQLKARMS	380
Db	1310	AALQEQPERRPSPHDEEFVEFEVAAEAQAQPKDGPSPAPASPEREVA-LSEYKTEYD	1368
Qy	381	KSXDGKTGSDD-----KKAKTSTRSS-AKTLKNRP	408
Db	1369	DYKDETTIDDSIMDADSLWVDTQDDRSIMTEQLETIPEKEAKEARRSLEKHREKP	1428
Qy	409	C-----LSPKLPTPGSSDPL-----IQPSSPA--VCPPEP	436
Db	1429	FKTGRGRISTPERKVAKEPSTVSRDEVRRRKAVYKKAELAKKTEVQAHSPSRKFIKPA	1488
Qy	437	---SSPKHVSVTSRPGSSGA-----KEMKLKGADGKTATPRGAAP-----PG	478
Db	1489	IKYTRTHLSVCKRKTATAGGESALAPSVFKQAKDKVSDGVTKSPKRSLLPRSSILPP	1548
Qy	479	QKG-----QANATRIPAKTP-----PAKPTPPSS-----GEPKPSG	509
Db	1549	RRGVSGDRDENSFSLNSSISSARRTTRSEPIRRAGKSGTSTPTPGSTAITPGTPPSYS	1608
Qy	510	DRSGYSRPGSTPGSRSTPSILPTTP-----TREPKVAVVTRTPKSPSSAKSRLQATP	564
Db	1609	SR-----TPGTGTPG-SYKPTPTGTPKSAILVPEKKKVAIIRTPPKSPATPK--OLRLIN	1662
Qy	565	VPMPDLKNVKSIGSTENILKHOPGGGKQIINKKLIDLNSVQSGKSGDNKHKVPGGVSQ	624
Db	1663	QLPDLKKNVKSIGSTDNIKYQKPGQVQIVTKIDLSH-----	1701
Qy	625	IVYKPVDLKSVTSKCSGLNIIHHKPGGGOVEVKSEKLDFFKORVOSKIGSLDNITHVPGGG	684
Db	1702	-----VTSKCSGLKNIRHREPGGRVKVTSYKCLDFEKAQAKVGSLDNAHHVPGGG	1751
Qy	685	NKKIETHKLTFFRENAKAKTDHCAETIVYKSPVVSVDGTSPRHLSNVSVSTSGSIDMVDSPOLAT	744
Db	1752	NVKIDSQKLNFRHAKARVYDHCAEIIITQSPGRSSVSPRSLNSVSSSGSINLLESQPOLAT	1811
Qy	745	LADEVASLAKOGL	758
Db	1812	LAEDVTAALAKOGL	1825

[illegible]

RESULT 11
US-09-724-676A-51331
; Sequence 51331, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD


```

Db 1702 -----VTSKCSGLNIRHPGGGRVKIESVKLDFEKAQAVGSLDNAHHVPGG 1751
QY 685 NKIETHKLFRENAKTDHGAIEIVYKSPVVSDDTSPRHLNVSTGSDMDVDSQLAT 744
Db 1752 NVKIDSKLNFREHAKARVDHGAIEIITQSPGRSVASPRRLSNVSSGSINLLESQLAT 1811
QY 745 LADEVASASLAKOGL 758
Db 1812 LAEDVTAALAKOGL 1825

RESULT 12
US-09-724-676-51332
; Sequence 51332, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 51332
; LENGTH: 471
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-51332

```

```

Query Match      18.9%; Score 750.5; DB 5; Length 471;
Best Local Similarity 38.3%; Pred. No. 1.1e-26;
Matches 199; Conservative 70; Mismatches 128; Indels 123; Gaps 18;

QY 297 APL-EFTFHVETPNVQKQAHSEHGLGAA--FP-----GAPGE-GPEARGPSLGEDT 346
Db 17 APLTEASASH--PPEIKDQGGAGELVRSANGFPYREDEGAFGEHGSQGTYSNTKENG 74
QY 347 KEADLPPEPEKQPAAPRGKPVSRVQLKARVYKSKDGTGDDKAKTSTRSSAKTLKN 406
Db 75 INGEL--TSADRETAEVRSARIVQVVTAEAVALK-----GQEKEAQHKDQTAAL---- 123
QY 407 RPCLSPKLPFTGSSDPLIQSPSPAVCPPEPPSPKHVSVTSRTGSSGA-----KEMK 458
Db 124 -----PLAAEETANLPPSPPPSPASEQTVTVEEAAGGESALAPSVFQKQAK 168
QY 459 LKGADGKTKIATPRGAAP-----PGQKG-----QANATRIPAKTP--- 493
Db 169 DKVSDGVTKSPKRSLSLPPRSSILPPRGVSGDRDENSEFSLNSSISSARRTRSEPIRR 228
QY 494 -----PAPKTPSS-----GEPKSGDRSGYSPGSPGTGSRSTRPSLTTPP-----TR 538
Db 229 AGKSGTSTPTPGSTAITPGTPPSYSSR---TPGTPGTP-SYPRTPHTPGTPKSAILVP 283
QY 539 EPKKVAVVTRTPPKSPSSAKSRLOTAPVPMPLKNVKSIGSTENLKHQPGGKGVQIINKK 598
Db 284 SEKKVAILIRTPPKSPATPK-QURLINQPLDPLKNVKSIGSTDNIKYQPKGGQVIVTKK 342
QY 599 LDLSNVQSKCGSKDNKIKHVPGGGSVQIVYKPVDSLKVTSKCSGLNIHHKPGGGOVEVKS 658
Db 343 IDLSH-----VTSCGSLKNIRHPPGGGRVKIES 371
QY 659 EKLDKDRVQSKIGSLDNIHVPGGGNKKIETHKLTFRENAKAKTDHGAIEIVYKSPVVS 718
Db 372 VKLDFEKAQAKVGLSDNAHHVPGGNGVKIDSQKLNFRHAKARVDHGAIEIITQSPGRSS 431
QY 719 DTSRHLNVSTGSDMDVDSQLATLADEVASASLAKOGL 758
Db 432 VASPRRLSNVSSGSINLLESQLATLAEDVTAALAKOGL 471

RESULT 13
US-09-724-676A-51332
; Sequence 51332, Application US/09724676A

```

```

; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 51332
; LENGTH: 471
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-51332

Query Match      18.9%; Score 750.5; DB 5; Length 471;
Best Local Similarity 38.3%; Pred. No. 1.1e-26;
Matches 199; Conservative 70; Mismatches 128; Indels 123; Gaps 18;

QY 297 APL-EFTFHVETPNVQKQAHSEHGLGAA--FP-----GAPGE-GPEARGPSLGEDT 346
Db 17 APLTEASASH--PPEIKDQGGAGELVRSANGFPYREDEGAFGEHGSQGTYSNTKENG 74
QY 347 KEADLPPEPEKQPAAPRGKPVSRVQLKARVYKSKDGTGDDKAKTSTRSSAKTLKN 406
Db 75 INGEL--TSADRETAEVRSARIVQVVTAEAVALK-----GQEKEAQHKDQTAAL---- 123
QY 407 RPCLSPKLPFTGSSDPLIQSPSPAVCPPEPPSPKHVSVTSRTGSSGA-----KEMK 458
Db 124 -----PLAAEETANLPPSPPPSPASEQTVTVEEAAGGESALAPSVFQKQAK 168
QY 459 LKGADGKTKIATPRGAAP-----PGQKG-----QANATRIPAKTP--- 493
Db 169 DKVSDGVTKSPKRSLSLPPRSSILPPRGVSGDRDENSEFSLNSSISSARRTRSEPIRR 228
QY 494 -----PAPKTPSS-----GEPKSGDRSGYSPGSPGTGSRSTRPSLTTPP-----TR 538
Db 229 AGKSGTSTPTPGSTAITPGTPPSYSSR---TPGTPGTP-SYPRTPHTPGTPKSAILVP 283
QY 539 EPKKVAVVTRTPPKSPSSAKSRLOTAPVPMPLKNVKSIGSTENLKHQPGGKGVQIINKK 598
Db 284 SEKKVAILIRTPPKSPATPK-QURLINQPLDPLKNVKSIGSTDNIKYQPKGGQVIVTKK 342
QY 599 LDLSNVQSKCGSKDNKIKHVPGGGSVQIVYKPVDSLKVTSKCSGLNIHHKPGGGOVEVKS 658
Db 343 IDLSH-----VTSCGSLKNIRHPPGGGRVKIES 371
QY 659 EKLDKDRVQSKIGSLDNIHVPGGGNKKIETHKLTFRENAKAKTDHGAIEIVYKSPVVS 718
Db 372 VKLDFEKAQAKVGLSDNAHHVPGGNGVKIDSQKLNFRHAKARVDHGAIEIITQSPGRSS 431
QY 719 DTSRHLNVSTGSDMDVDSQLATLADEVASASLAKOGL 758
Db 432 VASPRRLSNVSSGSINLLESQLATLAEDVTAALAKOGL 471

RESULT 14
US-10-070-611-1
; Sequence 1, Application US/10070611
; GENERAL INFORMATION:
; APPLICANT: VON BERGEN, Martin
; APPLICANT: BIERNAT, Jacek
; APPLICANT: MANDELKOW, Eva-Maria
; APPLICANT: MANDELKOW, Eckhard
; TITLE OF INVENTION: A MINIMAL TAU PEPTIDE FOR THE NUCLEATION OF PAIRED HELICAL FRA
; FILE REFERENCE: 029976/0101
; CURRENT APPLICATION NUMBER: US/10/070,611
; CURRENT FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: PCT/EP00/08863
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: EP 99117805.4
; PRIOR FILING DATE: 1999-09-09
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn version 3.1

```

; SEQ ID NO 1
; LENGTH: 125
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-070-611-1

Query Match 16.6%; Score 657; DB 6; Length 125;
Best Local Similarity 100.0%; Pred. No. 4.4e-23;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 561 QTAPVMPDLKNNVSKIGSTENLKHOPGGKVOIINKKLDLSNVQSKGSKDNIKHVPGG 620
Db 1 QTAPVMPDLKNNVSKIGSTENLKHOPGGKVOIINKKLDLSNVQSKGSKDNIKHVPGG 60
QY 621 GSVQIVYKPVLDLKVTSKSGSLGNIHHKPGGGQVEVSEKLFKDRVQSKIGSLDNITHV 680
Db 61 GSVQIVYKPVLDLKVTSKSGSLGNIHHKPGGGQVEVSEKLFKDRVQSKIGSLDNITHV 120

QY 681 PGGGN 685
Db 121 PGGGN 125

RESULT 15

US-10-100-957A-152
; Sequence 152, Application US/10100957A
; GENERAL INFORMATION:
; APPLICANT: Kapur, Ravi
; TITLE OF INVENTION: A System for Cell Based Screening
; FILE REFERENCE: 97-022-L1A
; CURRENT APPLICATION NUMBER: US/10/100,957A
; CURRENT FILING DATE: 2002-03-19
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 152
; LENGTH: 1125
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-100-957A-152

Query Match 15.7%; Score 621.5; DB 6; Length 1125;
Best Local Similarity 28.5%; Pred. No. 1.6e-20;
Matches 234; Conservative 91; Mismatches 246; Indels 251; Gaps 33;

QY 43 LKESPLQTPEDGSG---EE--PGSETSDAKSTPTAEDVTAPLVDEGAPGKQAAQPHTE 96
Db 317 LPTEPDESSTKDVAPPMEEEIVPCNNTTSPKETET---TLPIKMDLAPEDVLLTKETE 372
QY 97 IPEGTTAAEAGIGDTPSLEDEAAGHVTOBPESGVVQVEGFLREPGPPGLSHQIMSGMP-- 154
Db 373 L-----APAKGMVSLSEIEALAKNDVRSAEI--PVAQETWVSEVVLATEVVLPSDPIT 426
QY 155 -----GAPLLPGPRENTQPGSTGPTEDGGRHAPPELLKHOLLGDLHQEGPPLKAG- 207
Db 427 TLTKDVTLPLEARPLVDMTPS--LEMTLKGKETAPPTETN--LG--MAKDMSPLESEV 482
QY 208 --GKE-----RPGSKEEVDEDRDVEDSSPDQSPPSK----- 236
Db 483 TLGKDVILPETKVAENNVNTPLSEEVTSVKDMSPSAETAPLAKNADLHSGTELIVDN 542
QY 237 -ASPAQD-----GR---PPQTAAREAT 254
Db 543 SMAPASDLALPLETKVATVPIKDKGTVOOTEKPREDSQLASMQHKGQSTVPCTASPE-- 600
QY 255 SIFGFPAGGAIPLPVDFLSKVST----EIPASEPDGSPVGRAGQDAPLETFHVEITFN 310
Db 601 --PVKAAEQMSTLPIDAPSPLENLEQKETPGSQSPSCSVSRQEEAKA--VGVTCN 654
QY 311 VQKEQAHSEHLCRAAFPAGPGGPEARG-----PSLGEDTKADLPEPSEKQPA--- 360
Db 655 DITTPPNKE-----PPPSPEKKAKPLATTQTPAKTSTSKAKTQTPSLPKQAPPT 703

QY 361 -----AAPRGKP-----VSRVPQLKARVMVSKSDGTGSD--DKKAKTS 396
Db 704 SGGLNKKPMSLASGVSVAAPHKRPAAATATATPSTLTPARDV-KPKPITEAKVAEKRTSPS 762
QY 397 TRSSAKTLKNRPCLSKPLTPGSSDPLIQ--PS--SPA--VCPEPPSS-----PKHVSSV 445
Db 763 KPSSAPALKPGPKTPTTVSKATSPSTLVSTGSPSRSPATTLPKRPTSLKTEGKPADVKRM 822
QY 446 TSRTGSSGAKEMKLGADGKTKIATPPGAAPPQKGOANATRIPAKTPPAPKTPPPSSGEP 505
Db 823 TAKSASADLSRSKTTSSASSVKRNTTPTGAAPP--AGMTSTRV-----KPMASAPSRSSG 873
QY 506 PKSGDRSOSYSSPGSPGTPGSRSTPSLPTPTPREPKKVAVVVRTPPKSPSSAKSLQTA 565
Db 874 ALSVDKK-----PT-----STKPSSSAPRVSRLATT-V 900
QY 566 PMPDLKNVSKIGSTENLKHOPGG----- 589
Db 901 SAPDLKSVRSKVGSTENIKHQPGGGRKAVEKKEBTAATAGKPEPNAVTKAAGSTIASAQKP 960
QY 590 --GKVOIINKKLDLSNVQSKGSKDNIKHVPGGGVSQIVYKPVLDLSKVTSCGSLGNIHH 647
Db 961 PAGKVQIVYKVKVSYSHIQSKVCVSKDNIKHVPCCGNVQIQNKVKVDISKVSSKCGSKANIKH 1020
QY 648 KPGGGQVEVSEKLFKDRVQSKIGSLDNITHVPGGNNKIE 689
Db 1021 KPGGGDVKIESOKLNFKEKAQKAVGSLDNVGHFPAGGAVKTE 1062

Search completed: January 28, 2003, 14:39:20
Job time : 30.1541 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:10:57 ; Search time 15.7356 Seconds
(without alignments)
1417.333 Million cell updates/sec

Title: US-09-904-987-5
Perfect score: 3965
Sequence: 1 MAEPQFEVMDHAGTYGL.....SPQLATLADEVSASLAKOGL 758

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep: *
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep: *
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep: *
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep: *
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep: *
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2116.5	53.4	441	4	US-08-244-603A-1
2	2077.5	52.4	441	4	US-08-913-915-5
3	1776.5	44.8	383	4	US-09-336-038-4
4	1590	40.1	352	2	US-08-726-306A-17
5	1590	40.1	391	2	US-08-244-951A-10
6	1590	40.1	391	2	US-08-389-011-23
7	1590	40.1	391	3	US-08-403-917A-23
8	1590	40.1	391	4	US-08-348-952A-23
9	1586	40.0	351	1	US-08-159-969-2
10	733	18.5	140	4	US-08-913-915-9
11	621.5	15.7	1125	4	US-09-513-783A-152
12	621.5	15.7	1610	4	US-09-513-783A-22
13	593	15.0	112	3	US-08-666-360-1
14	575	14.5	109	4	US-08-913-915-3
15	563	14.2	106	3	US-08-776-404B-1
16	561	14.1	108	4	US-08-913-915-2
17	546.5	13.8	109	4	US-08-913-915-1
18	546	13.8	140	4	US-08-913-915-10
19	416	10.5	160	2	US-08-726-306A-35
20	364	9.2	67	2	US-08-389-011-1
21	364	9.2	67	3	US-08-403-917A-1
22	364	9.2	67	4	US-09-348-952A-1
23	339	8.5	67	2	US-08-244-951A-1
24	283.5	7.2	1057	3	US-08-931-820-4
25	283	7.1	1185	4	US-09-041-886-23
26	280.5	7.1	1442	2	US-08-316-650-12
27	280.5	7.1	1442	5	PCT-US95-02251-12

28 278.5 7.0 1078 3 US-08-963-825-21 Sequence 21, Appl
29 278.5 7.0 1078 4 US-09-500-811-21 Sequence 21, Appl
30 278.5 7.0 1078 4 US-09-570-573-21 Sequence 21, Appl
31 278.5 7.0 1078 4 US-09-548-608-21 Sequence 21, Appl
32 274 6.9 941 4 US-07-757-022B-14 Sequence 84, Appl
33 274 6.9 1022 4 US-07-757-022B-84 Sequence 74, Appl
34 274 6.9 1038 4 US-07-757-022B-74 Sequence 58, Appl
35 274 6.9 1049 4 US-07-757-022B-58 Sequence 104, App
36 274 6.9 1140 4 US-07-757-022B-104 Sequence 44, Appl
37 274 6.9 1270 4 US-07-757-022B-42 Sequence 42, Appl
38 274 6.9 1311 4 US-07-757-022B-142 Sequence 142, App
39 274 6.9 1313 4 US-07-757-022B-50 Sequence 50, Appl
40 274 6.9 1314 4 US-07-757-022B-46 Sequence 46, Appl
41 274 6.9 1320 4 US-07-757-022B-60 Sequence 60, Appl
42 274 6.9 1320 4 US-07-757-022B-48 Sequence 48, Appl
43 274 6.9 1354 4 US-07-757-022B-40 Sequence 40, Appl
44 274 6.9 1361 4 US-07-757-022B-40 Sequence 52, Appl
45 274 6.9 1363 4 US-07-757-022B-52

ALIGNMENTS

RESULT 1
US-08-244-603A-1
; Sequence 1, Application US/08244603A
; Patent No. 6200768
; GENERAL INFORMATION:
; APPLICANT: Mandelkow, Eva-Maria
; APPLICANT: Mandelkow, Eckhard
; APPLICANT: Lichtenberg-Kraag, Birgit
; APPLICANT: Biernat, Jacek
; APPLICANT: Drewes, Gerard
; APPLICANT: Steiner, Barbara
; TITLE OF INVENTION: No. 6200768el Tools For The Diagnosis And
; TITLE OF INVENTION: Treatment Of Alzheimer's Disease
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Borun
; STREET: 233 South Waker Drive, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Tape
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/244,603A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Joseph A. Williams, Jr.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 28384/32778
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-484-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 441 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-244-603A-1

Query Match 53.4%; Score 2116.5; DB 4; Length 441;
Best Local Similarity 58.2%; Pred. No. 1.3e-133;

		LENGTH: 441 amino acids		Query Match 52.4%; Score 2077.5; DB 4; Length 441;			
		TYPE: amino acid		Best Local Similarity 57.5%; Pred. No. 5.1e-131;			
		TOPOLOGY: linear		Matches 436; Conservative 0; Mismatches 5; Indels 317; Gaps 2;			
		MOLECULE TYPE: protein					
		US-08-913-915-5					
QY	1	MAEPROEFVEFHAGTYGLDRKDOGGYTMHQDQEGDTAGLKESPLQPTPTEDGSEEPG	60	QY	1	MAEPROEFVEFHAGTYGLDRKDOGGYTMHQDQEGDTAGLKESPLQPTPTEDGSEEPG	60
DB	1	MAEPROEFVEFHAGTYGLDRKDOGGYTMHQDQEGDTAGLKESPLQPTPTEDGSEEPG	60	DB	1	MAEPROEFVEFHAGTYGLDRKDOGGYTMHQDQEGDTAGLKESPLQPTPTEDGSEEPG	60
QY	61	SETSDAKSTPTAEDVTAPLVDEGAPGKQAAQPHTEIPGTTAEAGIGDTPSLEDAAG	120	QY	61	SETSDAKSTPTAEDVTAPLVDEGAPGKQAAQPHTEIPGTTAEAGIGDTPSLEDAAG	120
DB	61	SETSDAKSTPTAEDVTAPLVDEGAPGKQAAQPHTEIPGTTAEAGIGDTPSLEDAAG	120	DB	61	SETSDAKSTPTAEDVTAPLVDEGAPGKQAAQPHTEIPGTTAEAGIGDTPSLEDAAG	120
QY	121	HVTQEPESGKVQEGFLREPGLSHQLMSGMPGAPLLPEGPREATRQPSGTGPEDTEG	180	QY	121	HVTQEPESGKVQEGFLREPGLSHQLMSGMPGAPLLPEGPREATRQPSGTGPEDTEG	180
DB	121	HVTQ-----	124	DB	121	HVTQ-----	124
QY	181	GRHAPELLKHQLLDGLHQEGPLKGAGKERPGSGKEEVEDRDVDDESSPDSPSKASPA	240	QY	181	GRHAPELLKHQLLDGLHQEGPLKGAGKERPGSGKEEVEDRDVDDESSPDSPSKASPA	240
DB	125	-----	124	DB	125	-----	124
QY	241	QDGRPPQTAAREATSIPGPAEGAIPLVDFLSKVSTEIPASEPDGSPVGRAGQDAPLE	300	QY	241	QDGRPPQTAAREATSIPGPAEGAIPLVDFLSKVSTEIPASEPDGSPVGRAGQDAPLE	300
DB	125	-----	124	DB	125	-----	124
QY	301	FTFHVETPNVQKEQAHSEHLGRAAPGAPGEGPEARGPSLGEDTKADLPPESEKQPA	360	QY	301	FTFHVETPNVQKEQAHSEHLGRAAPGAPGEGPEARGPSLGEDTKADLPPESEKQPA	360
DB	125	-----	124	DB	125	-----	124
QY	361	AAPRGKPVSRVQPKARMVSKSDGTGSDDKAKTSTRSSAKTLKNRCLSPKLPPTGSS	420	QY	361	AAPRGKPVSRVQPKARMVSKSDGTGSDDKAKTSTRSSAKTLKNRCLSPKLPPTGSS	420
DB	125	-----	142	DB	125	-----	142
QY	421	DPLIQSSPAVCPEPPSPKHVSSVTSRTGSSGAKEMKLKADGKTKIATPRGAAPPQK	480	QY	421	DPLIQSSPAVCPEPPSPKHVSSVTSRTGSSGAKEMKLKADGKTKIATPRGAAPPQK	480
DB	143	-----	163	DB	143	-----	163
QY	481	GOANATRIAPKTPAPKTPPSSGEPKSGDRSGYSSPGSPGTPGSRSTPSLPTPTREP	540	QY	481	GOANATRIAPKTPAPKTPPSSGEPKSGDRSGYSSPGSPGTPGSRSTPSLPTPTREP	540
DB	164	GOANATRIAPKTPAPKTPPSSGEPKSGDRSGYSSPGSPGTPGSRSTPSLPTPTREP	223	DB	164	GOANATRIAPKTPAPKTPPSSGEPKSGDRSGYSSPGSPGTPGSRSTPSLPTPTREP	223
QY	541	KKVAVVRTPKSPSSAKSRLOTAPVMPDLKNVKSIGSTENLKHQPGGKQVQIINKKLD	600	QY	541	KKVAVVRTPKSPSSAKSRLOTAPVMPDLKNVKSIGSTENLKHQPGGKQVQIINKKLD	600
DB	224	KKVAVVRTPKSPSSAKSRLOTAPVMPDLKNVKSIGSTENLKHQPGGKQVQIINKKLD	283	DB	224	KKVAVVRTPKSPSSAKSRLOTAPVMPDLKNVKSIGSTENLKHQPGGKQVQIINKKLD	283
QY	601	LSNVQSKCGSKDNKIKHVPGGSVQIVYKPVDSLKVTSKCSLGNIIHHKPGGQVEVKSEK	660	QY	601	LSNVQSKCGSKDNKIKHVPGGSVQIVYKPVDSLKVTSKCSLGNIIHHKPGGQVEVKSEK	660
DB	284	LSNVQSKCGSKDNKIKQVPGGSGVQIVYKPVDSLKVTSKCSLGNIIHHKPGGQVEVKSEK	343	DB	284	LSNVQSKCGSKDNKIKQVPGGSGVQIVYKPVDSLKVTSKCSLGNIIHHKPGGQVEVKSEK	343
QY	661	LDKDRVQSKIGSLDNIITHVPGGKTKIETHKLTFRENAKAKTDHGAETIVYKSPVVS	720	QY	661	LDKDRVQSKIGSLDNIITHVPGGKTKIETHKLTFRENAKAKTDHGAETIVYKSPVVS	720
DB	344	LDKDRVQSKIGSLDNIITHVPGGKTKIETHKLTFRENAKAKTDHGAETIVYKSPVVS	403	DB	344	LDKDRVQSKIGSLDNIITHVPGGKTKIETHKLTFRENAKAKTDHGAETIVYKSPVVS	403
QY	721	SPRHLSNVSTGSDIMVDSPOLATLADEVSAKQGL	758	QY	721	SPRHLSNVSTGSDIMVDSPOLATLADEVSAKQGL	758
DB	404	SPRHLSNVSTGSDIMVDSPOLATLADEVSAKQGL	441	DB	404	SPRHLSNVSTGSDIMVDSPOLATLADEVSAKQGL	441
RESULT 3		US-09-336-038-4					
		Sequence 4, Application US/09336038					
		Patent No. 6417185					
		GENERAL INFORMATION:					
		APPLICANT: Goff, Dane					
		APPLICANT: Harrison, Steven					
		APPLICANT: Nuss, John					
		APPLICANT: Ring, David B.					
		APPLICANT: Zhou, Xiaohui A.					

```

; TITLE OF INVENTION: Inhibitors of Glycogen Synthase Kinase 3
; FILE REFERENCE: 1443.003
; CURRENT APPLICATION NUMBER: US/09/336,038
; CURRENT FILING DATE: 1999-06-18
; EARLIER APPLICATION NUMBER: 60/089,978
; EARLIER FILING DATE: 1998-06-19
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-336-038-4

```

Query Match	44.8%;	Score	1776.5;	DB	4;	Length	383;
Best Local Similarity	50.5%;	Pred. No.	4.9e-111;				
Matches	383;	Conservative	0;	Mismatches	0;	Indels	375;
Gaps	3;						
QY	1	MAEPROEFVEMDHAGT	YGLGDRKDQGGY	TMHQDEGDT	DAGLKESPL	QTPTEDGSEEP	60
Db	1	MAEPROEFVEMDHAGT	YGLGDRKDQGGY	TMHQDEGDT	DAGLK		44
QY	61	SETSDAKSTPTAEDVT	APLVDEGAPGQAAAQ	HPTEIGECTTAAE	AGIGDTPSLEDA	EAAG	120
Db	45						
QY	121	HVTQEPESKVVQEGFL	REBPGLSLHQLMSG	MPGAPLLPEGP	REATRQPSGTGP	DEDTG	180
Db	63	HVTQ					66
QY	181	GRHAPELLKHLLGDL	HQHEGPP	PLKGAGGKERPGS	KEEVDREDVD	ESSQDPSPPSK	240
Db	67						66
QY	241	QDGRPQTAAREATS	IPGFPAGAIPL	VPDLFSKVSTEI	PASEPDGPSV	GRAKGQDAP	300
Db	67						66
QY	301	FTFHVEITPNVQKEQ	AHSEHLGRAAP	GAPGEGPEARG	PSLGEDTK	KEADLPEPSEK	360
Db	67						66
QY	361	AAPRGKPVSRVPQL	KARMVSKSDGT	QSDDKKAKT	STRSSAKTL	KNRPLSPKLP	420
Db	67						84
QY	421	DPLIQSSPAVCP	PEPPSPKHSV	SVTSRTGSGAGK	EMKLKGADGKY	KIATPRGAAP	480
Db	85						105
QY	481	GOANATRIPAKTP	PAPKTPPSGSEPP	KSGDRSGYS	SPSGPTPGSR	SRTPSLPTPT	540
Db	106	GOANATRIPAKTP	PAPKTPPSGSEPP	KSGDRSGYS	SPSGPTPGSR	SRTPSLPTPT	165
QY	541	KKVAVVRTPPKSP	SSAKSRLQ	TAPVMPDL	KNVKSKITG	STENLKHQGGK	600
Db	166	KKVAVVRTPPKSP	SSAKSRLQ	TAPVMPDL	KNVKSKITG	STENLKHQGGK	225
QY	601	LSNVQSKCGSKNI	KHVPGGGS	QVIYKVPD	LSKVTCKG	SLGNIHHKPGG	660
Db	226	LSNVQSKCGSKNI	KHVPGGGS	QVIYKVPD	LSKVTCKG	SLGNIHHKPGG	285
QY	661	LDFKRVQSKIGSL	DNITHVPGGN	KNKTETHK	LTFRENAK	AKTDHGAIEI	720
Db	286	LDFKRVQSKIGSL	DNITHVPGGN	KNKTETHK	LTFRENAK	AKTDHGAIEI	345
QY	721	SPRHLNSVSTGS	IDMWDS	PQ	LADI	EV	758
Db	346	SPRHLNSVSTGS	IDMWDS	PQ	LADI	EV	383

RESULT 4
US-08-726-306A-17

```

; Sequence 17, Application US/08726306A
; Patent No. 5958684
; GENERAL INFORMATION:
; APPLICANT: van Leeuwen, Frederik Willem
; APPLICANT: Burbach, Johannes Peter Henri
; APPLICANT: Grosveld, Franklin G.
; TITLE OF INVENTION: DIAGNOSIS METHOD AND REAGENTS
; NUMBER OF SEQUENCES: 189
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Witcoff, Ltd.
; STREET: 1 Financial Center
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Wordperfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/726,306A
; FILING DATE: 02-Oct-1996
; PRIOR APPLICATION DATA:
; FILING APPLICATION NUMBER: GB 95/20080.4
; FILING DATE: 02-Oct-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/009,832
; FILING DATE: 01-Jan-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams, Ph.D., Kathleen M.
; REGISTRATION NUMBER: 34,380
; REFERENCE/DOCKET NUMBER: 96,048-A (3255/00784)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 345-9100
; TELEFAX: (617) 345-9111
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 352 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-726-306A-17

```

```

Query Match      40.1%; Score 1590; DB 2; Length 352;
Best Local Similarity 46.4%; Pred. No. 1.2e-98;
Matches 352; Conservative 0; Mismatches 0; Indels 406; Gaps 4;

QY 1 MAERQREFEWMEDHAGNYGLGDRKDQGGYTMHQDQEGDTDAGLKESPLQPTPDGSEEPG 60
      |||||
Db 1 MAERQREFEWMEDHAGNYGLGDRKDQGGYTMHQDQEGDTDAGLK----- 44
      |||||

QY 61 SETSDAKSTPTAEDVTAPLVDGAPGQAQAAPHTPEIGTAEAGIGDTPTPSLEDEAAG 120
      |||||
Db 45 -----AAEAGIGDTPTPSLEDEAAG 62
      |||||

QY 121 HVTQEPESGKVVGEGFLREPGPGLSHQLMSGMPGAPLLPEGPREATRQPSGTGPDTEG 180
      ||||
Db 63 HVTQ----- 66

QY 181 GRHAPELLKHQLLDLHQEGPPLKAGGKERPGSKKEEVEDEDRDVEDSSPQDSPPSKASPA 240
      67 ----- 66

QY 241 QDGRPPQTAAREATSI PGFPAEGAIPLPVDFLSKVSTEIPASEPDGPSVGRAKQDAPLE 300
      67 ----- 66

QY 301 FTFHVEITPNVQKEQAHSSEHLGRAAPGAPGEGPEARGPSLGEDTKEADLPPESEKOPA 360
      67 ----- 66

```

QY 361 AAPRGKPVSRVQQLKARVSKSKDGTGSDDKKAKTSTRSSAKTLKNRPCLSPKLPPTPGSS 420
Db 67 -----ARMVSKSKDGTGSDDKA----- 84
QY 421 DPLIQSPSPAVCPPEPPSPKHSVSVTSRTGSSGAKEMKLGADGKTKIATPRCAAPPGOK 480
Db 85 -----KGADGKTKIATPRCAAPPGOK 105
QY 481 GOANATRIAPAKTPAPKTPSSGEPKSGDRSGYSSPGSPGTPGSRSTRPSLPTPTREP 540
Db 106 GOANATRIAPAKTPAPKTPSSGEPKSGDRSGYSSPGSPGTPGSRSTRPSLPTPTREP 165
QY 541 KKVAVVTPPKSPSSAKSRQLQTAAPVMPDLKNVSKIGSTENLKHQPGGGKVOIINKKLD 600
Db 166 KKVAVVTPPKSPSSAKSRQLQTAAPVMPDLKNVSKIGSTENLKHQPGGGK----- 216
QY 601 LSNVQSKGSKDNKHKVPGGGSVOIVYKPDLSKVTSKGSLGNIHHKPGGGQVEVKSEK 660
Db 217 -----VOIVYKPDLSKVTSKGSLGNIHHKPGGGQVEVKSEK 254
QY 661 LDFKDRVQSKIGSLDNITHVPGGKNKIETHKLTFRENAKAKTDHGAIEIVYKSPVVSQGT 720
Db 255 LDFKDRVQSKIGSLDNITHVPGGKNKIETHKLTFRENAKAKTDHGAIEIVYKSPVVSQGT 314
QY 721 SPRHLSNVSTGSDIMVDSPOLATLADEVASLAKQGL 758
Db 315 SPRHLSNVSTGSDIMVDSPOLATLADEVASLAKQGL 352

RESULT 5

US-08-244-951A-10
; Sequence 10, Application US/08244951A
; Patent No. 5843779

GENERAL INFORMATION:

; APPLICANT: VANDERMEEREN, MARC; MERCKEN, MARC;
; APPLICANT: VANMECHELEN, EUGEN; VAN DE VOORDE, ANDRE
; TITLE OF INVENTION: MONOCLONAL ANTIBODIES
; TITLE OF INVENTION: DIRECTED AGAINST THE MICROTUBULE-ASSOCIATED
; TITLE OF INVENTION: PROTEIN TAU, HYBRIDOMAS SECRETING THESE
; TITLE OF INVENTION: ANTIBODIES, ANTIGEN RECOGNITION BY THESE
; TITLE OF INVENTION: MONOCLONAL ANTIBODIES AND THEIR APPLICATIONS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: BIERMAN & MUSERLIAN
; STREET: 600 THIRD AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10016

COMPUTER READABLE FORM:

; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII

CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/244, 951A
; FILING DATE: 19-JAN-1995

CLASSIFICATION:

; CLASSIFICATION: 435

PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/EP93/03499

FILING DATE:

; FILING DATE: 10-DEC-1993

PRIOR APPLICATION DATA:

; APPLICATION NUMBER: EP/92/403403.6

FILING DATE:

; FILING DATE: 14-DEC-1992

ATTORNEY/AGENT INFORMATION:

; NAME: CHARLES A. MUSERLIAN

REGISTRATION NUMBER:

; REGISTRATION NUMBER: 19,683

REFERENCE/DOCKET NUMBER:

; REFERENCE/DOCKET NUMBER: 410.003A

TELEPHONE:

; TELEPHONE: (212) 661-8000

TELEFAX:

; TELEFAX: (212) 661-8002

INFORMATION FOR SEQ ID NO:

; INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

; LENGTH: 391
; TYPE: Amino Acid
; STRANDEDNESS: Unknown
; TOPOLOGY: Unknown
; FEATURE:
; NAME/KEY: mTHMPH-taul fusion protein
; US-08-244-951A-10

Query Match

Best Local Similarity 40.1%; Score 1590; DB 2; Length 391;

Matches 352; Conservative 46.4%; Pred. No. 1.4e-98;

Mismatches 0; Indels 406; Gaps 4;

QY 1 MAEPQREFEVMEHDHAGTYGLGRKDDQGGYTMHQDQEGDTHAGLAKESPLQPTEDGSEBP 60

Db 40 MAEPQREFEVMEHDHAGTYGLGRKDDQGGYTMHQDQEGDTHAGLAKESPLQPTEDGSEBP 60

QY 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAQPHITEPEGTTAEAGIGDTPSLEDAAG 120

Db 84 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAQPHITEPEGTTAEAGIGDTPSLEDAAG 120

QY 121 HVTQEPESGVVQEGFLREPGLSHQMSGMPGAPLLPEGPREATRQPSGTGPEDETEG 180

Db 102 HVTQEPESGVVQEGFLREPGLSHQMSGMPGAPLLPEGPREATRQPSGTGPEDETEG 180

QY 181 GRHAPELLKHQLLDHQLHOGEPPLKAGGCKERPSSKEEVEDRDVDDESSPDSPSKASPA 240

Db 106 GRHAPELLKHQLLDHQLHOGEPPLKAGGCKERPSSKEEVEDRDVDDESSPDSPSKASPA 240

QY 241 QDGRPPQTAAREATSIPIGFPAGGAIPVDFLSKVSTEIPASEPDGSPVGRAKQDAPLE 300

Db 106 QDGRPPQTAAREATSIPIGFPAGGAIPVDFLSKVSTEIPASEPDGSPVGRAKQDAPLE 300

QY 301 FTFHVEITPNVQEQAHSEHLGRAAPGAPGEGPEARGPSLGEDTKEADLPEPSEKQPA 360

Db 106 FTFHVEITPNVQEQAHSEHLGRAAPGAPGEGPEARGPSLGEDTKEADLPEPSEKQPA 360

QY 361 AAPRGKPVSRVQQLKARVSKSKDGTGSDDKKAKTSTRSSAKTLKNRPCLSPKLPPTPGSS 420

Db 106 AAPRGKPVSRVQQLKARVSKSKDGTGSDDKKAKTSTRSSAKTLKNRPCLSPKLPPTPGSS 420

QY 421 DPLIQSPSPAVCPPEPPSPKHSVSVTSRTGSSGAKEMKLGADGKTKIATPRGAAPPGOK 480

Db 124 DPLIQSPSPAVCPPEPPSPKHSVSVTSRTGSSGAKEMKLGADGKTKIATPRGAAPPGOK 480

QY 481 GOANATRIAPAKTPAPKTPSSGEPKSGDRSGYSSPGSPGTPGSRSTRPSLPTPTREP 540

Db 145 GOANATRIAPAKTPAPKTPSSGEPKSGDRSGYSSPGSPGTPGSRSTRPSLPTPTREP 540

QY 541 KKVAVVTPPKSPSSAKSRQLQTAAPVMPDLKNVSKIGSTENLKHQPGGGKVOIINKKLD 600

Db 205 KKVAVVTPPKSPSSAKSRQLQTAAPVMPDLKNVSKIGSTENLKHQPGGGKVOIINKKLD 600

QY 601 LSNVQSKGSKDNKHKVPGGGSVOIVYKPDLSKVTSKGSLGNIHHKPGGGQVEVKSEK 660

Db 256 LSNVQSKGSKDNKHKVPGGGSVOIVYKPDLSKVTSKGSLGNIHHKPGGGQVEVKSEK 660

QY 661 LDFKDRVQSKIGSLDNITHVPGGKNKIETHKLTFRENAKAKTDHGAIEIVYKSPVVSQGT 720

Db 294 LDFKDRVQSKIGSLDNITHVPGGKNKIETHKLTFRENAKAKTDHGAIEIVYKSPVVSQGT 720

QY 721 SPRHLSNVSTGSDIMVDSPOLATLADEVASLAKQGL 758

Db 354 SPRHLSNVSTGSDIMVDSPOLATLADEVASLAKQGL 758

RESULT 6

US-08-389-011-23
; Sequence 23, Application US/08389011
; Patent No. 5861257

GENERAL INFORMATION:

; APPLICANT: VANDERMEEREN, MARC; MERCKEN, MARC;
; APPLICANT: VANMECHELEN, EUGEN; VAN DE VOORDE, ANDRE
; TITLE OF INVENTION: MONOCLONAL ANTIBODIES

TITLE OF INVENTION: DIRECTED AGAINST THE MICROTUBULE-ASSOCIATED
 TITLE OF INVENTION: PROTEIN TAU, HYBRIDOMAS SECRETING THESE
 TITLE OF INVENTION: ANTIBODIES, ANTIGEN RECOGNITION BY THESE
 TITLE OF INVENTION: MONOCLONAL ANTIBODIES AND THEIR APPLICATIONS.
 NUMBER OF SEQUENCES: 24
 CORRESPONDENCE ADDRESS:

ADDRESSEE: BIERMAN & MUSERLIAN
 STREET: 600 THIRD AVENUE
 CITY: NEW YORK
 STATE: NEW YORK
 COUNTRY: USA
 ZIP: 10016

COMPUTER READABLE FORM:
 MEDIUM TYPE: FLOPPY DISK
 COMPUTER: IBM PC COMPATIBLE
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: ASCII

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/389,011
 FILING DATE: 15-FEB-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/403,917
 FILING DATE: 19-JAN-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/403,916
 FILING DATE: 19-JAN-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/244,951
 FILING DATE: 13-JUN-1994

APPLICATION DATA:
 APPLICATION NUMBER: PCT/EP93/03499
 FILING DATE: 10-DEC-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: EP/92/403403.6
 FILING DATE: 14-DEC-1992

ATTORNEY/AGENT INFORMATION:
 NAME: CHARLES A. MUSERLIAN
 REGISTRATION NUMBER: 19,683
 REFERENCE/DOCKET NUMBER: 410.003-1-CON
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 661-8000
 TELEFAX: (212) 661-8002

INFORMATION FOR SEQ ID NO: 23:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 391
 TYPE: Amino Acid
 STRANDEDNESS: Unknown
 TOPOLOGY: Unknown

US-08-389-011-23

Query Match 40.1%; Score 1590; DB 2; Length 391;
 Best Local Similarity 46.4%; Pred No. 1.4e-98;
 Matches 352; Conservative 0; Mismatches 0; Indels 406; Gaps 4;

QY 1 MAEPROEFVEMDHAGTYGLGDRKQGGYTMHQDQEGDTAGLKESPLQTPTEDGSEEPG 60
 DB 40 MAEPROEFVEMDHAGTYGLGDRKQGGYTMHQDQEGDTAGLKESPLQTPTEDGSEEPG 60
 QY 61 SETSDAKSTPTAEDYAPLVDEGAPGKQNAAPHTPEIGETTAEBAGIGDTPSLDEAAG 120
 DB 84
 QY 121 HVTQPESGKVQVGEFLRPPGPPGLSHQLMSOMGAPLLPEGPREATQPSGTGPDTEG 180
 DB 102 HVTQ
 QY 181 GRHAPELLKHQLLDHQLHOGPPPLKGAGKERPGSKEEVEDRDVDESSPQSPSKASPA 240
 DB 106
 QY 241 QDGRPPQTAAREATSIPIGPAEATPLPVDFLSKVSTEIPASEPDGSPVGRAGKQDAPLE 300

DB 106
 QY 301 FTFHVEITPNVQEQAHSEHGLGRAAFPAGFEGPEARGLSGLDTEADLPSEKQPA 360
 DB 106
 QY 361 AAPRGKPVSRVQPLKARMVSKSDGTGSDDKAKKTSTRSSAKTLKNRCLSPKLPPTPGSS 420
 DB 106
 QY 421 DPLIQSPSPAVCPPEPPSPKHVSSVTSRTSGSGAKEMKLGADGKTATPRGAAPGQK 480
 DB 124
 QY 481 QOANATRIPAKTPPAKTPPSSGEPKSGDRSGYSSPGSPCTPGSRSTRPSLPTPTREP 540
 DB 145 QOANATRIPAKTPPAKTPPSSGEPKSGDRSGYSSPGSPCTPGSRSTRPSLPTPTREP 204
 QY 541 KKVAVVRTPPKSPSSAKSRLQAPVMPDLKNVSKIGSTENLKHQPGGKGVQIINKLD 600
 DB 205 KKVAVVRTPPKSPSSAKSRLQAPVMPDLKNVSKIGSTENLKHQPGGKGVQIINKLD 255
 QY 601 LSNVQSKCGSKDNKHPVGGGVQIVYKPVDSLKVTSCGSLGNHHPKGGGQVEVKSEK 660
 DB 256
 QY 661 LDFKDRVQSKIGSLDNITHVPGGNNKKTETHKLTFRENAKAKTDHGAIVYKSPVWSGDT 720
 DB 294 LDFKDRVQSKIGSLDNITHVPGGNNKKTETHKLTFRENAKAKTDHGAIVYKSPVWSGDT 353
 QY 721 SPRHLSNVSTGSDIMVDPOLATLADSVASLAKOGL 758
 DB 354 SPRHLSNVSTGSDIMVDPOLATLADSVASLAKOGL 391

RESULT 7

US-08-403-917A-23
 ; Sequence 23 Application US/08403917A
 ; Patent No. 6010913

GENERAL INFORMATION:
 APPLICANT: VANDEMEEREN, MARC; MERCKEN, MARC;
 APPLICANT: VANMECHELEN, EUGEN;

APPLICANT: VAN DE VOORDE, ANDRE
 TITLE OF INVENTION: MONOCLONAL ANTIBODIES

TITLE OF INVENTION: DIRECTED AGAINST THE MICROTUBULE-ASSOCIATED
 TITLE OF INVENTION: PROTEIN TAU, HYBRIDOMAS SECRETING THESE
 TITLE OF INVENTION: ANTIBODIES, ANTIGEN RECOGNITION BY THESE
 TITLE OF INVENTION: MONOCLONAL ANTIBODIES AND THEIR APPLICATION
 NUMBER OF SEQUENCES: 24

CORRESPONDENCE ADDRESS:

ADDRESSEE: BIERMAN & MUSERLIAN
 STREET: 600 THIRD AVENUE
 CITY: NEW YORK
 STATE: NEW YORK
 COUNTRY: USA
 ZIP: 10016

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK

COMPUTER: IBM PC COMPATIBLE

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/403,917A

FILING DATE: 19-JAN-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/256,167

FILING DATE: 27-JUN-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/244,951

FILING DATE: 13-JUN-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/EP93/03499

FILING DATE: 10-DEC-1993

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: EP/92/403403.6
;; FILING DATE: 14-DEC-1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: CHARLES A. MUSERLIAN
;; REGISTRATION NUMBER: 19,683
;; REFERENCE/DOCKET NUMBER: 410.003-1
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212) 661-8000
;; TELEFAX: (212) 661-8002
;; INFORMATION FOR SEQ ID NO: 23:
;; SEQUENCE CHARACTERISTICS:
;; TYPE: Amino Acid
;; STRANDEDNESS: Unknown
;; TOPOLOGY: Unknown
;;
US-08-403-917A-23

Query Match 40.1%; Score 1590; DB 3; Length 391;
Best Local Similarity 46.4%; Pred. No. 1.4e-98;
Matches 352; Conservative 0; Mismatches 0; Indels 406; Gaps 4;

Qy 1 MAEPQEFVEMDHAGTYGLGDRKQGGYTMHQDQGGDTDAGLKESPLQTPTEDESGSEPG 60
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||| 83
Qy 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAQPHTEIPGTTAEAGIGDTPSLEDEAAG 120
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||| 101
Qy 121 HVTQEPESGVVQEGFLREPGLSHQLMSGMPGAPLLPEGPREATRQPSGTGPEDTEG 180
Db |||| 105
Qy 102 HVTQ- 105
Qy 181 GRHAPELLKHQLLDLHQEGPPLKAGGKERPGSKEEVEDRDVDDESSQDPSPKASPA 240
Db ----- 105
Qy 241 QDGRPPQTAAREATSIPGFFAEGAIPLPVDFLSKVSTEIPASEPDGPDGVSRAKQDAPLE 300
Db ----- 105
Qy 301 FTFHVETPNVQREAHSEHLGRAAPGAPGEGPEARGPSLGEDTKEADLPEPSEKQPA 360
Db ----- 105
Qy 361 AAPRGKPVSRVPQKLKRWVSKSDGTGSDDKKAKTSTRSSAKTLKNRPCLSPKLPTPGSS 420
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||| 123
Qy 421 DPLIQSSPAVCPPEPPSSPKHVSVTSRTGSSGAKEMKLGADGKTKIATPRGAAPPQOK 480
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||| 144
Qy 481 QGANATRIKATPPAKTPPSPGEPKSGDRSGYSSPGSPGTPGSRSTRPSLPTPTREP 540
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||| 204
Qy 541 KKVAVTTPKSPSSAKSRLOTAPVMPDLKNVSKSIGSTENLKHQPGGKGVQIINKLKD 600
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||| 255
Qy 601 LSNVQSCGSKDNKHPVGGSGVQIVKPYDLKSVTSKCSGLNIHKKPGGGQVEVKSEK 660
Db -----VQIVKPYDLKSVTSKCSGLNIHKKPGGGQVEVKSEK 293
Qy 661 LDFKDRVQSKIGSLDNTHVPGGKNKIETHKLTFRENAKAKTDHGAIEIVKSPVVSQGT 720
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||| 353
Qy 721 SPRHLNSVSTGSDMVDSPQLATLADSVASLAKQGL 758
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||| 391
Qy 354 SPRHLNSVSTGSDMVDSPQLATLADSVASLAKQGL 391

RESULT 8
US-09-348-952A-23
; Sequence 23, Application US/09348952A
; Patent No. 6232437
; GENERAL INFORMATION:
; APPLICANT: VANDERMEEREN, MARC; MERCKEN, MARC;
; APPLICANT: VANMECHELEN, EUGEN;
; APPLICANT: VAN DE VOORDE, ANDRE
; TITLE OF INVENTION: MONOCLONAL ANTIBODIES
; TITLE OF INVENTION: DIRECTED AGAINST THE MICROTUBULE-ASSOCIATED
; TITLE OF INVENTION: PROTEIN TAU, HYBRIDOMAS SECRETING THESE
; TITLE OF INVENTION: ANTIBODIES, ANTIGEN RECOGNITION BY THESE
; TITLE OF INVENTION: MONOCLONAL ANTIBODIES AND THEIR APPLICATION
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIERMAN & MUSERLIAN
; STREET: 600 THIRD AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/348,952A
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/403,917
; FILING DATE: 19-JAN-1995
; APPLICATION NUMBER: 08/256,167
; FILING DATE: 27-JUN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/244,951
; FILING DATE: 13-JUN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP93/03499
; FILING DATE: 10-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP/92/403403.6
; FILING DATE: 14-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: CHARLES A. MUSERLIAN
; REGISTRATION NUMBER: 19,683
; REFERENCE/DOCKET NUMBER: 410.003-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 661-8000
; TELEFAX: (212) 661-8002
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 391
; TYPE: Amino Acid
; STRANDEDNESS: Unknown
; TOPOLOGY: Unknown
; US-09-348-952A-23

Query Match 40.1%; Score 1590; DB 4; Length 391;
Best Local Similarity 46.4%; Pred. No. 1.4e-98;
Matches 352; Conservative 0; Mismatches 0; Indels 406; Gaps 4;

Qy 1 MAEPQEFVEMDHAGTYGLGDRKQGGYTMHQDQGGDTDAGLKESPLQTPTEDESGSEPG 60
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||| 83
Qy 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAQPHTEIPGTTAEAGIGDTPSLEDEAAG 120
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||| 101
Qy 121 HVTQEPESGVVQEGFLREPGLSHQLMSGMPGAPLLPEGPREATRQPSGTGPEDTEG 180


```

;
; NAME: Misrock, S.Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 4697-040
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 351 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-159-969-2

```

Query Match 40.0%; Score 1586; DB 1; Length 351;

Best Local Similarity 46.4%; Pred. No. 2.2e-98;

Matches 351; Conservative 0; Mismatches 0; Indels 406; Gaps 4;

```

QY 1 MAEPROEEFVEMDHAGTYGLGDRKDOGGYTMHQDQEGDTHDAGLKESPLQPTPTDSEEPG 60
DB 1 MAEPROEEFVEMDHAGTYGLGDRKDOGGYTMHQDQEGDTHDAGLKESPLQPTPTDSEEPG 44
QY 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAQPHTEIPEGTTAEAGIGDTTPSLEDEAAG 120
DB 45 -----AEEAGIGDTTPSLEDEAAG 62
QY 121 HVTQEPESKVVQEGFLREPGLSHQLMSGMPGAPLLPEGPREATROPSTGPTDTEG 180
DB 63 HVTQ-----AEEAGIGDTTPSLEDEAAG 66
QY 181 GRHAPELLKHQLLDLHQEGPPLKGAGKERPGSKKEEVEDRDVDSESPDPSKASPA 240
DB 67 -----AEEAGIGDTTPSLEDEAAG 66
QY 241 QDGRPPQTAAREATSIPGPPAEGAIPLVPDLSKVSTEIPASEPDGPGSVGRAKQDAPLE 300
DB 67 -----AEEAGIGDTTPSLEDEAAG 66
QY 301 FTFHVEITPNVQEQAHSEHLGRAAFPAGCEGPEARGPSLGDTKEADLPPESEKQPA 360
DB 67 -----AEEAGIGDTTPSLEDEAAG 66
QY 361 AAPRGKPVSRVQLKARWVSKSDGTGDDKAKTSTRSSAKTLKNRPCLSPKLPTPGSS 420
DB 67 -----AEEAGIGDTTPSLEDEAAG 84
QY 421 DPLIQSPSPAVCPPEPPSPKHVSVTSRTGSSGAKEMKLKGADGKTATPRGAAPPQK 480
DB 85 -----KGADGKTATPRGAAPPQK 105
QY 481 GOANATRIPAKTPAPKTPPSSGEPKSGDRSGYSSPGSPGTPGSRSTPSLPTPTREP 540
DB 106 GOANATRIPAKTPAPKTPPSSGEPKSGDRSGYSSPGSPGTPGSRSTPSLPTPTREP 165
QY 541 KKVAVVTRTPPKSPSSAKSRLQAPVMPDLKNVSKIGSTENLKHQPGGKQVLIINKLKD 600
DB 166 KKVAVVTRTPPKSPSSAKSRLQAPVMPDLKNVSKIGSTENLKHQPGGKQVLIINKLKD 216
QY 601 LSNVQSKCGSKDNIRKHPVGGGSVQIVYKPVDSLKVTSKCGSLGNTHHKPGGQGVVEKSEK 660
DB 217 -----QIVYKPVDSLKVTSKCGSLGNTHHKPGGQGVVEKSEK 254
QY 661 LDFKDRVQSKIGSLDNITHVPGGNNKKTETHKLTFRENAKAKTDHGAETVYKSPVVSQDT 720
DB 255 LDFKDRVQSKIGSLDNITHVPGGNNKKTETHKLTFRENAKAKTDHGAETVYKSPVVSQDT 314
QY 721 SPRHLNSVSTGSDIMVDSQPLATLADEVASLAKQGL 757
DB 315 SPRHLNSVSTGSDIMVDSQPLATLADEVASLAKQGL 351

```

RESULT 10

```

;
; US-08-159-969-2
; Sequence 2, Application us/08159969
; Patent No. 5492812
; GENERAL INFORMATION:
; APPLICANT: Voorhees, Paul H.
; TITLE OF INVENTION: Diagnostic Method for Alzheimer's
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/159,969
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/738,778
; FILING DATE: 01-AUG-1991
; ATTORNEY/AGENT INFORMATION:

```

US-09-913-915-9

; Sequence 9, Application US/08913915

; Patent No. 6376205

; GENERAL INFORMATION:

; APPLICANT:

; TITLE OF INVENTION: INHIBITION OF TAU-TAU-ASSOCIATION

; NUMBER OF SEQUENCES: 12

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/913,915

; FILING DATE:

; INFORMATION FOR SEQ ID NO: 9:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 140 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; US-08-913-915-9

Query Match 18.5%; Score 733; DB 4; Length 140;

Best Local Similarity 100.0%; Pred. No. 4.8e-42;

Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 569 DLKNNKIGSTENLKHQPGGKVIINKLIDLSNVQSKGSKDNKIHVPGGSGVQIVYK 628

DB 1 DLKNNKIGSTENLKHQPGGKVIINKLIDLSNVQSKGSKDNKIHVPGGSGVQIVYK 60

QY 629 PVDLSKVTSCGSLGNTHHKPGGQVEVKESEKLDKDRVQSKIGSLDNIHVPGGNGKKI 688

DB 61 PVDLSKVTSCGSLGNTHHKPGGQVEVKESEKLDKDRVQSKIGSLDNIHVPGGNGKKI 120

QY 689 ETHKLTFRNKAKTDHGAE 708

DB 121 ETHKLTFRNKAKTDHGAE 140

RESULT 11

US-09-513-783A-152

; Sequence 152, Application US/09513783A

; Patent No. 6416959

; GENERAL INFORMATION:

; APPLICANT: Giuliano, Kenneth A.

; TITLE OF INVENTION: A System for Cell Based Screening

; FILE REFERENCE: 97-022-L1

; CURRENT APPLICATION NUMBER: US/09/513,783A

; CURRENT FILING DATE: 2000-02-25

; NUMBER OF SEQ ID NOS: 180

; SOFTWARE: Patentin Ver. 2.0

; SEQ ID NO 152

; LENGTH: 1125

; TYPE: PRT

; ORGANISM: Mus musculus

US-09-513-783A-152

Query Match

Best Local Similarity 28.5%; Score 621.5; DB 4; Length 1125;

Matches 234; Conservative 91; Mismatches 246; Indels 251; Gaps 33;

QY 43 LKESPLQTPTEDGS-----EE--PCEYSDAKSTPTAEDVTAPLYDEGAPGKQAAQPHTE 96

DB 317 LPTEPDESSTKDVAPPMEETVPGNDITSPKETET---TLPIKMDLAPPEDVLLTKETE 372

QY 97 IPEGTTAEAGIGDTPTSLDEAAGHVTOEPESGKVGQEGFLRPGPGSLHQLMSGMP-- 154

DB 373 L-----APAKGMVLSLSEEEALAKNDVRSAL-I-PVAQETVVSETEVVLATVLPSPDIT 426

QY 155 -----GAPLLPEGPREATRQPSGTGPTDEGGRHAPPELLKHQLLDLHOGEPPLKAG- 207

DB 427 TLTQDVTLPLEAERPLVTOMTPS-LETEMTLGKETAPPTETN--LC-MAKDMSPLESEV 482

QY 208 --GKE-----RPGSKEEVEDRDVDESSPODPPSK----- 236

DB 483 TLGQDVVILPETKVAEFNNVTPLSSEEVTSVKDMSPSAETEAFLAKNADLHSGTGLIVDN 542

QY 237 -ASPAOD-----PPSPPEKKAKPLATTQPAKTSTSKAKTQPTSLPKQAPPT 703

DB 543 SWAPASDLALPLETKVATVPIKDKGTQVTEKEPREDSQLASMQHKGQSTVPPTCTASPE-- 600

QY 255 SIPIGPPAEGAILPLVDLFSKVT-----EIPASEPDPGSPVGRAKGQDAPLEFFTHVEITPN 310

DB 601 --PVKAAEQMSTLPIDAPSPLENLEQETPGSQSPSPSPATTLKRPRTSIKTEGKPADVKRM 822

QY 311 VQKEQAHSEHLGRAAFPAGGEPARG-----PSLGEDTKADLPEPESEKOPA--- 360

DB 655 DITTPPNKE-----PPSPPEKKAKPLATTQPAKTSTSKAKTQPTSLPKQAPPT 703

QY 361 -----AAPRGKP-----VSRVPQLKARMVSKSKDGTGSD--DKKAKTS 396

DB 704 SGLNKKPMKSLASGSPVPAAPHKRPAAATATARPSTLPARDV-KPKPITEAKVAEKRTSPS 762

QY 397 TRSSAKTLNRCPLSKLPTPGSSDPLIQ--PS--SPA-VCPEPPSS-----PKHVSIV 445

DB 763 KPSSAPALKPGPKTPTVSKATSPSTLVSTGSPSRSPATTLKRPRTSIKTEGKPADVKRM 822

QY 446 TSRTGSGGAKEMKLGADCKTKIATPRGAAPGQKQOANATRIPAKTPAPKTPPSSGEP 505

DB 823 TAKSADISLRSKTTSSASSVKRNTTPTGAAP--AGMTSTRV-----KPMASPSRSG 873

QY 506 PKSGDRSGYSSGSPGTPGSRSTRPSLPTPTREPKKVAVVRTPPKSPSSAKSLQITAPV 565

DB 874 ALSVDKK-----PT-----STKPPSSAPRYSRLATT-V 900

QY 566 PMPDLKNVSKIGSTENLKHQPGG----- 589

DB 901 SAPDLKSVRSKVGSTENIKHQPGGRKRAKVEKTEAATTAGKPEPNNAVTKAAGSIASAKP 960

QY 590 --GKVOIINKKLDLSNVQSKGSKDNKIHVPGGSGVQIVYKVDLSKVTSCGSLGNTIHH 647

DB 961 PAGQVIVSKVSYSHIQSKCVSKDNKIHVPGGNGVQIONKKVDSKSSKCGSKANIKH 1020

QY 648 KPGGGQVEVKESEKLDKDRVQSKIGSLDNIHVPGGNGKKIE 689

DB 1021 KPGGGDVKIESQKLNPKERAKQAKVGSGLDNVGHFPAAGGAVKTE 1062

RESULT 12

US-09-513-783A-22

; Sequence 22, Application US/09513783A

; Patent No. 6416959

; GENERAL INFORMATION:

; APPLICANT: Giuliano, Kenneth A.

; APPLICANT: Kapor, Ravi

; TITLE OF INVENTION: A System for Cell Based Screening

; FILE REFERENCE: 97-022-L1

; CURRENT APPLICATION NUMBER: US/09/513,783A

; CURRENT FILING DATE: 2000-02-25

; NUMBER OF SEQ ID NOS: 180

; SOFTWARE: Patentin Ver. 2.0

; SEQ ID NO 22

; LENGTH: 1610

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence:

; OTHER INFORMATION: EYFP-DEVD-MAP4-EBFP Construct

US-09-513-783A-22

Query Match

Best Local Similarity 15.7%; Score 621.5; DB 4; Length 1610;

Matches 234; Conservative 91; Mismatches 246; Indels 251; Gaps 33;


```

1  TITLE OF INVENTION: ANTIBODIES AND THEIR APPLICATIONS
2
3  NUMBER OF SEQUENCES: 5
4
5  CORRESPONDENCE ADDRESS:
6
7  ADDRESSEE: A. ROLD, WHITE & DURKEE
8
9  STREET: P.O. BOX 4433
10
11 CITY: HOUSTON
12
13 STATE: TEXAS
14
15 COUNTRY: USA
16
17 ZIP: 77210-4433
18
19 COMPUTER READABLE FORM:
20
21 MEDIUM TYPE: Floppy disk
22
23 COMPUTER: IBM PC compatible
24
25 OPERATING SYSTEM: PC-DOS/MS-DOS
26
27 SOFTWARE: Microsoft Word 6.0 / ASCII text output
28
29 CURRENT APPLICATION DATA:
30
31 APPLICATION NUMBER: US/08/776,404B
32
33 FILING DATE: 27 Jan 1997
34
35 PRIOR APPLICATION DATA:
36
37 APPLICATION NUMBER: PCT/EP95/03032
38
39 FILING DATE: 31 Jul 1995
40
41 PRIOR APPLICATION DATA: EP 94870131.3
42
43 APPLICATION NUMBER: EP 94870131.3
44
45 FILING DATE: 29 Jul 1994
46
47 ATTORNEY/AGENT INFORMATION:
48
49 NAME: KAMMERER, PATRICIA A.
50
51 REGISTRATION NUMBER: 29,775
52
53 REFERENCE/DOCKET NUMBER: INNS:003
54
55 INFORMATION FOR SEQ ID NO: 1:
56
57 SEQUENCE CHARACTERISTICS:
58
59 LENGTH: 106 amino acids
60
61 TYPE: amino acid
62
63 STRANDEDNESS: single
64
65 TOPOLOGY: linear
66
67 MOLECULE TYPE: peptide
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
97
```

```
Query Match      14.2%; Score 563; DB 3; Length 106;
Best Local Similarity 100.0%; Pred. No. 7.3e-31;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 463 DGKTKIATPRGAAPPQKGQANATRIPAKTPPSSGEPKSGDRSGSPGT 522

Db 1 DGKTKIATPRGAAPPGOKGKANATRIPAKTPPPAPKTPPPSSGEPKSGDRSCYSSPGSPCT 60

QY 523 PGSRRTPSLTPTPTREPKKVAVVRTPPKSPSSAKSRLQTA PVMP 568

Db 61 PGSRSRTPSLPTTREPCKKVAVRTPPKSPSSAKSRLTAPVPMP 106

Search completed: January 28, 2003, 14:21:40
Job time : 33.7356 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:08:34 ; Search time 4.03824 Seconds
(without alignments)
3666.126 Million cell updates/sec

Title: US-09-904-987-6

Perfect score: 821

Sequence: 1 MATKAVCVLKGDPVQGIN.....STKTGNAGSLACGVIGIAQ 154

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	821	100.0	154	1 DSHUCZ	superoxide dismuta
2	807	98.3	179	4 T43640	superoxide dismuta
3	684.5	83.4	152	2 S36108	superoxide dismuta
4	680	82.8	154	2 J00915	superoxide dismuta
5	678	82.6	154	2 J01192	superoxide dismuta
6	667	81.2	154	1 DSHOCZ	superoxide dismuta
7	659.5	80.3	153	2 S33162	superoxide dismuta
8	652	79.4	152	1 DSHOCZ	superoxide dismuta
9	650.5	78.2	152	1 DSHOCZ	superoxide dismuta
10	635	77.3	151	2 A24475	superoxide dismuta
11	599	73.0	153	2 S65436	superoxide dismuta
12	589.5	71.8	166	2 S29782	superoxide dismuta
13	558.5	68.0	152	2 S04623	superoxide dismuta
14	536.5	65.3	151	1 DSHOCZ	superoxide dismuta
15	533	64.9	151	1 S05568	superoxide dismuta
16	521	63.5	150	2 S05021	superoxide dismuta
17	512	62.4	153	2 A45171	superoxide dismuta
18	505	61.5	153	2 S48117	superoxide dismuta
19	493	60.0	153	1 DSHOCZ	superoxide dismuta
20	493	60.0	153	2 S05498	superoxide dismuta
21	482	59.9	153	2 A49241	superoxide dismuta
22	482	59.3	153	2 S03606	superoxide dismuta
23	482.5	58.8	154	2 S20511	superoxide dismuta
24	480	58.5	151	2 A29077	superoxide dismuta
25	477.5	58.2	152	2 S21136	superoxide dismuta
26	469.5	57.2	152	2 S07007	superoxide dismuta
27	465.5	56.7	201	2 T06800	superoxide dismuta
28	464.5	56.6	216	2 T51730	superoxide dismuta
29	464.5	56.6	222	1 DSSPCZ	superoxide dismuta

30	463.5	56.5	202	1 DSPMCZ	superoxide dismuta
31	463.5	56.5	211	2 T03685	probable superoxid
32	462.5	56.3	152	2 S22508	superoxide dismuta
33	462.5	56.3	201	2 T06229	probable superoxid
34	462	56.3	158	2 S18743	superoxide dismuta
35	461.5	56.2	152	2 S72235	superoxide dismuta
36	460	56.0	158	2 A48256	superoxide dismuta
37	455.5	55.5	156	2 T12204	superoxide dismuta
38	455.5	55.5	219	2 S03608	superoxide dismuta
39	453.5	55.2	217	2 S48021	superoxide dismuta
40	451.5	55.0	152	2 T06570	superoxide dismuta
41	451.5	55.0	162	2 T51731	superoxide dismuta
42	450.5	54.9	218	2 H84681	superoxide dismuta
43	443.5	54.0	152	2 S40404	probable copper/z1
44	441.5	53.8	152	2 JW0084	superoxide dismuta
45	439.5	53.5	152	1 DSSPCY	superoxide dismuta

ALIGNMENTS

RESULT 1

DSHUCZ

superoxide dismutase (EC 1.15.1.1) (Cu-Zn) [validated] - human

C:Species: Homo sapiens (man)

C:Date: 31-Aug-1980 #sequence_revision 19-Feb-1984 #text_change 21-Jul-2000

C:Accession: A22703; PC2217; A23046; A93969; JX0055; A91282; A90440; T54382; I68676;

R:Levanon, D.; Lieman-Hurwitz, J.; Dafni, N.; Wigderson, M.; Sherman, L.; Bernstein,

EMBO J. 4, 77-84, 1985

A:Title: Architecture and anatomy of the chromosomal locus in human chromosome 21 enc

A:Reference number: A22703; MUID:85257452; PMID:3160582

A:Accession: A22703

A:Molecule type: DNA

A:Residues: 2-154 <LEV>

A:CROSS-references: EMBL:X01780; NID:g36534; PIDN:CAA25915.1; PID:g36535; EMBL:X01781

A:Note: the authors did not translate the codon for residue 1

R:Kim, H.T.; Kim, Y.H.; Nam, J.W.; Lee, H.J.; Rho, H.M.; Jung, G.

Biochem. Biophys. Res. Commun. 201, 1526-1533, 1994

A:Title: Study of 5'-flanking region of human Cu/Zn superoxide dismutase.

A:Reference number: PC2217; MUID:94296434; PMID:8024598

A:Accession: PC2217

A:Molecule type: DNA

A:Residues: 1-24 <KIM>

A:CROSS-references: EMBL:Z29336; NID:g531099

R:Hallewell, R.A.; Maslary, F.R.; Najarian, R.C.; Puma, J.P.; Quiroga, M.R.; Randolph

Nucleic Acids Res. 13, 2017-2034, 1985

A:Title: Human Cu/Zn superoxide dismutase cDNA: isolation of clones synthesising high

A:Reference number: A23046; MUID:85215596; PMID:3889846

A:Accession: A23046

A:Molecule type: mRNA

A:Residues: 1-154 <HAL>

A:CROSS-references: EMBL:X02317; NID:g36541; PIDN:CAA26182.1; PID:g36542

R:Sherman, L.; Dafni, N.; Lieman-Hurwitz, J.; Groner, Y.

Proc. Natl. Acad. Sci. U.S.A. 80, 5465-5469, 1983

A:Title: Nucleotide sequence and expression of human chromosome 21-encoded superoxide

A:Reference number: A93969; MUID:83299994; PMID:6577438

A:Accession: A93969

A:Molecule type: mRNA

A:Residues: 2-154 <SHE>

A:CROSS-references: GB:X00065; NID:g36541; PIDN:CAA26182.1; PID:g36542

A:Note: the authors did not translate the codon for residue 1

R:Kajihara, J.; Enomoto, M.; Nishijima, K.; Yabuuchi, M.; Katoh, K.

J. Biochem. 104, 851-854, 1988

A:Title: Comparison of properties between human recombinant and placental copper-zinc

A:Reference number: JX0055; MUID:89174523; PMID:2853161

A:Accession: JX0055

A:Molecule type: protein

A:Residues: 2-154 <RAJ>

A:Experimental source: placenta

R:Barra, D.; Martini, F.; Bannister, J.V.; Schinina, M.E.; Rotilio, G.; Bannister, W.

FEBS Lett. 120, 53-56, 1980

A:Title: The complete amino acid sequence of human Cu/Zn superoxide dismutase.

A:Reference number: A91282; MUID:81067132; PMID:7002610

A:Accession: A91282
 A:Molecule type: protein
 A:Residues: 2-11,'N',13-92,'N',94-154 <BAR>
 A:Experimental source: erythrocytes
 R:Jabusch, J.R.; Farb, D.L.; Kerschensteiner, D.A.; Deutsch, H.F.
 Biochemistry 19, 2310-2316, 1980
 A:Title: Some sulfhydryl properties and primary structure of human erythrocyte superoxide
 A:Reference number: A90440; MUID:80221052; PMID:6770891
 A:Accession: A90440
 A:Molecule type: protein
 A:Residues: 2-17,'S',19-26,'N',28-49,'Q',51-52,'ND',55-98,'V',100-154 <TAB>
 R:Enayati, Z.E.; Orrell, R.W.; Claus, A.; Ludolph, A.; Bachus, R.; Brockmueller, J.; Ray-
 cell, J.
 Hum. Mol. Genet. 4, 1239-1240, 1995
 A:Title: Two novel mutations in the gene for copper zinc superoxide dismutase in UK fami
 A:Reference number: I54382; MUID:96133303; PMID:8528216
 A:Accession: I54382
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 'SMKKQIH',127-154 <ENAL>
 A:Cross-references: GB:L44746; NID:G928824; PIDN:AAC41773.1; PID:G928825
 A:Note: mutated exon 5 from patient with familial amyotrophic lateral sclerosis
 A:Accession: I68676
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 25-48,'Q',50-56 <ENAZ>
 A:Cross-references: GB:L46374; NID:G939772; PIDN:AAB59626.1; PID:G939773
 A:Note: mutant sequence from patient with familial amyotrophic lateral sclerosis
 A:Accession: I68677
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 120-125,'H',127-154 <ENAZ>
 A:Cross-references: GB:L46375; NID:G939774; PIDN:AAB59627.1; PID:G939775
 A:Note: mutant sequence from patient with familial amyotrophic lateral sclerosis
 R:Yamazaki, Y.; Takao, T.; Murata, H.; Kawaharada, Y.; Sugiyama, T.; Shimonishi, Y.
 Res. Commun. Biochem. Cell Mol. Biol. 1, 205-217, 1997
 A:Title: The use of capillary liquid chromatography/electrospray mass spectrometry to id
 A:Reference number: JC5686
 A:Accession: JC5686
 A:Molecule type: protein
 A:Residues: 2-154 <YAM>
 A:Note: Inter-molecular disulfide bonds form under non-physiological conditions
 R:Rosen, D.R.; Siddique, T.; Patterson, D.; Figlewicz, D.A.; Sapp, P.; Hentati, A.; Dona
 n, S.M.; Berger, R.; Tanzi, R.E.; Halperin, J.J.; Herzfeldt, B.; Van den Bergh, R.; Hung
 Rouleau, G.A.; Gusella, J.S.; Horvitz, H.R.; Brown Jr., R.H.
 Nature 362, 59-62, 1993
 A:Title: Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyot
 A:Reference number: A44737; MUID:93188938; PMID:8446170
 A:Contents: annotation; mutant sequences reported
 R:Pargie, H.E.; Tainer, J.A.
 submitted to the Brookhaven Protein Data Bank, July 1993
 A:Reference number: A52244; PDB:1SPD
 A:Contents: annotation; X-ray crystallography, 2.4 angstroms, residues 2-154
 A:Note: erythrocyte recombinant form expressed in yeast
 R:Pargie, H.E.; Hallewell, R.A.; Tainer, J.A.
 Proc. Natl. Acad. Sci. U.S.A. 89, 6109-6113, 1992
 A:Title: Atomic structures of wild-type and thermostable mutant recombinant human Cu,Zn
 A:Reference number: A44540; MUID:92335247; PMID:1463506
 A:Contents: annotation; X-ray crystallography, 2.5 angstroms
 C:Genetics:
 A:Gene: GDB:SOD1; ALS; ALS1
 A:Cross-references: GDB:119596; OMIM:147450
 A:Map position: 21q22.1-21q22.1
 A:Introns: 24/3; 56/3; 80/3; 119/3
 A:Note: a defect in this enzyme is the cause of adult familial amyotrophic lateral scler
 C:Complex: homodimer
 C:Function:
 A:Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxygen
 C:Superfamily: superoxide dismutase (Cu-Zn)
 C:Keywords: acetylated amino end; copper; homodimer; Lou Gehrig's disease; metalloprotei
 F:2-154/Product: superoxide dismutase (Cu-Zn) #status experimental <MAT>
 F:2/Modified site: acetylated amino end (Ala) (in mature form) #status experimental

F:47,49,64,121/Binding site: copper (His) #status experimental
 F:58-147/Disulfide bonds: #status experimental
 F:64,72,81,84/Binding site: zinc (His, His, His, Asp) #status experimental
 F:144/Active site: Arg #status experimental
 Query Match 100.0%; Score 821; DB 1; Length 154;
 Best Local Similarity 100.0%; Pred. No. 1.6e-69;
 Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVHEFGDNTAGCTS 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVHEFGDNTAGCTS 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 120
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 120
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 RESULT 2
 T43640
 superoxide dismutase (EC 1.15.1.1) (Cu-Zn) - synthetic (fragment)
 C:Species: synthetic
 A:Note: human SOD1 and SOD3 genes engineered and expressed in Schizosaccharomyces pom
 C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 08-Dec-2000
 C:Accession: T43640
 R:Inoue, M.; Watanabe, N.; Matsuno, K.; Sasaki, J.; Tanaka, Y.; Hatanaka-Sawai, H.; A
 J. Biol. Chem. 266, 16409-16414, 1991
 A:Title: Expression of a hybrid Cu/Zn-type superoxide dismutase which has high affinity
 A:Reference number: Z22590; MUID:91358420; PMID:1885572
 A:Accession: T43640
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-179 <INO>
 A:Cross-references: EMBL:D90358; NID:G218563; PIDN:BAAL4373.1; PID:G218564
 A:Experimental source: strain EH13-15
 A:Note: the engineered initiator codon and the termination codon are not shown
 C:Comment: For the original sequences, see PIR:DSHUCZ and PIR:DSHUEC.
 C:Keywords: copper; metalloprotein; oxidoreductase; zinc
 F:1-151/Region: SOD1 gene derived
 F:153-179/Region: SOD3 gene derived
 F:46,48,63,120/Binding site: copper (His) #status predicted
 F:57-146/Disulfide bonds: #status predicted
 F:63,71,80,83/Binding site: zinc (His, His, His, Asp) #status predicted
 F:143/Active site: Arg #status predicted

Query Match 98.3%; Score 807; DB 4; Length 179;
 Best Local Similarity 100.0%; Pred. No. 3.8e-68;
 Matches 151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 ATKAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVHEFGDNTAGCTS 61
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1 ATKAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVHEFGDNTAGCTS 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 62 GPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 121
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 61 GPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 120
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 122 EKADDLGKGGNEESTKTGNAGSRLACGVIGI 152
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 121 EKADDLGKGGNEESTKTGNAGSRLACGVIGI 151
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 3

S36108

Superoxide dismutase (EC 1.15.1.1) (Cu-Zn) - guinea pig
 C:Species: Cavia porcellus (guinea pig)
 C:Date: 19-Mar-1997 #sequence_revision 09-Aug-1997 #text_change 20-Apr-2000
 C:Accession: S36108; S55796
 R:Wolf, B.; Reinecke, K.; Amann, K.D.; Brigelius-Flohe, R.; Flohe, L.

Biol. Chem. Hoppe-Seyler 374, 641-649, 1993
A:Title: Taxonomical classification of the guinea pig based on its Cu/Zn superoxide dismutase
A:Reference number: S36108; MUID:94059414; PMID:8240718
A:Accession: S36108
A:Molecule type: protein
A:Residues: 1-101, 114, 104-152 <WOL>
R:Yuan, H.T.; Bingle, C.D.; Kelly, F.J.
Biochem. Biophys. Acta 1305, 163-171, 1996
A:Title: Differential patterns of antioxidant enzyme mRNA expression in guinea pig lung
A:Reference number: S65793; MUID:96180320; PMID:8597602
A:Accession: S65796
A:Molecule type: mRNA
A:Residues: 6-145 <YUA>
A:Cross-references: EMBL:U39844
C:Function:
A:Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxygen
C:Superfamily: superoxide dismutase (Cu-Zn)
C:Keywords: copper; metalloprotein; oxidoreductase; zinc
F:45,47,62,119/Binding site: copper (His) #status predicted
F:56-145/Disulfide bonds: #status predicted
F:62,70,79,82/Binding site: zinc (His, His, Asp) #status predicted
F:142/Active site: Arg #status predicted

Query Match 83.4%; Score 684.5; DB 2; Length 152;
Best Local Similarity 86.3%; Pred. No. 8.6e-57;
Matches 132; Conservative 8; Mismatches 12; Indels 1; Gaps 1;

QY 2 ATRACVCLKGDGPVQGIINFEQKESNPVKVMSIKGLTEGLHGFHVFEGDNTAGCTSA 61
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 ATRACVCLKGDGPVQGIINFEQK-ANGPVVVKGRITGLVEGKHGFHVFEGDNTAGCTSA 59
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 62 GPHFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSIEDSVISLGDHCHIGRTLVVH 121
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 60 GPHFNPLSRKHGPKDEERHVGDLGNVTAGDGVANVSIEDSVISLGSANSIIGRTVMVH 119
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 122 EKADDLKGGNEESTKTGNAGSLACGVIGIAQ 154
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 120 EKPDLLKGGNEESTKTGNAGSLACGVIGIAQ 152
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 4
JQ0915
superoxide dismutase (EC 1.15.1.1) (Cu-Zn) - mouse
N:Alternate names: proteasome chain C8
C:Species: Mus musculus (house mouse)
C:Date: 30-Sep-1991 #sequence revision 30-Sep-1991 #text_change 20-Apr-2000
C:Accession: JQ0915; S00910; A37101
R:Benedetto, M.T.; Anzai, Y.; Gordon, J.W.
Gene 99, 191-195, 1991
A:Title: Isolation and analysis of the mouse genomic sequence encoding Cu2+-Zn2+ superoxide dismutase
A:Reference number: JQ0915; MUID:91216458; PMID:2022332
A:Accession: JQ0915
A:Molecule type: DNA
A:Residues: 1-154 <BEN>
A:Cross-references: GB:M60794
A:Note: the authors translated the codon CAT for residue 102 as Asp
A:Note: the sequence shown follows the authors' translation
R:Bewley, G.C.
Nucleic Acids Res. 16, 2728, 1988
A:Title: cDNA and deduced amino acid sequence of murine Cu-Zn superoxide dismutase.
A:Reference number: S00910; MUID:88203220; PMID:3362683
A:Accession: S00910
A:Molecule type: mRNA
A:Residues: 1-154 <BEW>
A:Cross-references: EMBL:X06683; NID:g54127; PIDN:CAA29880.1; PID:g54128
R:Pluchero, F.G.; Shreeve, M.; Eskinazi, D.; van der Gaag, H.; Huang, K.S.; Hulmes, J.D.
J. Cell Biol. 111, 1217-1223, 1990
A:Title: Purification of an inhibitor of erythroid progenitor cell cycling and antagonism
A:Reference number: A37101; MUID:90361747; PMID:2391363
A:Accession: A37101
A:Molecule type: protein
A:Residues: 4-6, 'X', 8-23 <PLU>

C:Genetics:

A:Gene: SOD-1

A:Introns: 24/3; 57/1; 80/2; 119/3

C:Function:

A:Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxygen
C:Superfamily: superoxide dismutase (Cu-Zn)
C:Keywords: copper; metalloprotein; oxidoreductase; zinc
F:47,49,64,121/Binding site: copper (His) #status predicted
F:58-147/Disulfide bonds: #status predicted
F:64,72,81,84/Binding site: zinc (His, His, Asp) #status predicted
F:144/Active site: Arg #status predicted

Query Match 82.8%; Score 680; DB 2; Length 154;

Best Local Similarity 83.8%; Pred. No. 2.3e-56;

Matches 129; Conservative 8; Mismatches 17; Indels 0; Gaps 0;

QY 1 MATKAVCVLKGDGPVQGIINFEQKESNPVKVMSIKGLTEGLHGFHVFEGDNTAGCTSA 60

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Db 1 MAMKAVCVLKGDGPVQGIINFEQKASGPVVLGGITGLTEGQHGHHVQYNDNTQGCTSA 60

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 61 AGPHFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSIEDSVISLGDHCHIGRTLVV 120

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Db 61 AGPHFNPLSRKHGPKDEERHVGDLGNVTAGDGVANVSIEDRVISLGSHEHSIIGRTMVV 120

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 121 HEKADDLKGGNEESTKTGNAGSLACGVIGIAQ 154

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Db 121 HEKQDDLKGGNEESTKTGNAGSLACGVIGIAQ 154

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 5

JC1192

superoxide dismutase (EC 1.15.1.1) (Cu-Zn) [validated] - rat

N:Alternate names: SOD Cu-Zn

C:Species: Rattus norvegicus (Norway rat)

C:Date: 05-Mar-1993 #sequence revision 05-Mar-1993 #text_change 20-Apr-2000

C:Accession: JC1192; A26810; A30595; A25570; JC2015; S00036; I53629

R:Hsu, J.L.; Visner, G.A.; Burr, I.A.; Nick, H.S.

Biochem. Biophys. Res. Commun. 186, 936-943, 1992

A:Title: Rat copper/zinc superoxide dismutase gene: Isolation, characterization, and
A:Reference number: JC1192; MUID:92360048; PMID:1379810

A:Accession: JC1192

A:Molecule type: DNA

A:Residues: 1-154 <HSU>

A:Cross-references: GB:S41731

R:Ho, Y.S.; Crapo, J.D.

Nucleic Acids Res. 15, 6746, 1987

A:Title: cDNA and deduced amino acid sequence of rat copper-zinc-containing superoxid

A:Reference number: A26810; MUID:87316943; PMID:3628012

A:Accession: A26810

A:Molecule type: mRNA

A:Residues: 1-154 <HOY>

A:Cross-references: GB:Y00404; GB:M31634; NID:g57274; PIDN:CAA68465.1; PID:g57275

R:Hass, M.A.; Iqbal, J.; Clerch, L.B.; Frank, L.; Massaro, D.

J. Clin. Invest. 83, 1241-1246, 1989

A:Title: Rat lung Cu,Zn superoxide dismutase. Isolation and sequence of a full-length

A:Reference number: A30595; MUID:89198078; PMID:2703531

A:Accession: A30595

A:Molecule type: mRNA

A:Residues: 1-154 <HAS>

A:Cross-references: GB:M25157

R:Steffens, G.J.; Michelson, A.M.; Puget, K.; Flohe, L.

Biol. Chem. Hoppe-Seyler 367, 1017-1024, 1986

A:Title: The amino-acid sequence of rat Cu-Zn superoxide dismutase.

A:Reference number: A25570; MUID:87076037; PMID:3790250

A:Accession: A25570

A:Molecule type: protein

A:Residues: 2-154 <STE>

R:Kim, Y.H.; Yoo, H.Y.; Jung, G.; Kim, J.Y.; Rho, H.M.

Gene 133, 267-271, 1993

A:Title: Isolation and analysis of the rat genomic sequence encoding Cu/Zn superoxide

A:Reference number: JC2015; MUID:94040822; PMID:8224914

A:Accession: JC2015

A:Molecule type: DNA

	Query Match	81.2%	Score 667;	DB 1:	Length 154;
	Best Local Similarity	81.0%;	Pred. No. 3.8e-55;		
	Matches 12;	Conservative 14;	Mismatches 15;	Indels 0;	Gaps 0;
Qy	1	MATKAVCLKGDGPVGIIINFEOKESNPVKVWGSTKGLTEGLUGHGFHVEFGDNTAGCTS	60		
Db	1	MALKAVCLKGDGPVGHVIFEQOEGGPVVLKGFTEGLTKDGHGFHVEFGDNTGGCTT	60		
Qy	61	AGPHFNLSRKHGPKDEERHVGDLGNVTKADKQGVADVSTEDSVISLSDGHCIIGRTLVV	120		
Db	61	AGAHFNLSRKHGPKDEERHVGDLGNVADENKADVDMDKDSVISLSGKHSITIGRTVW	120		

```
C:Species: Oryzotrogus concoloratus (domestic rabbit)
C:Date: 03-Mar-1994 #sequence_revision 10-Nov-1995 #text_change 20-Apr-2000
C:Accession: S33162; S01134
R:Jackson, R.M.; Ho, Y.
```

A;Description: Nucleotide and deduced amino acid sequence of rabbit CuZnSD.
A;Reference number: S33162
A;Accession: S33162

A;Residues: 1-153 <JAC>
A;Cross-references: EMBL:226644; NID:g296917; PTDN:CAA80357.1; PID:g296918
R;Reinscke, K.; Wolf, B.; Michelson, A.M.; Puget, K.; Steffens, G.J.; Flohe,

A;Title: The amino-acid sequence of rabbit Cu-Zn superoxide dismutase.
A;Reference number: S01134; MUID:89105317; PMID:3214553
A;Accession: S01134

A:Residues: 2-16, 'G',18-43, 'G',45-53, 'T',55-151, 'A',153 <REI>
 A:Note: the sequence from Fig. 5 is inconsistent with that from Fig. 4 in having 132-
 C:Function:

A: description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxygen
C: Superfamily: superoxide dismutase (Cu-Zn)
E: 46.48.63.120/Binding site: copper (His) #status predicted
P: 46.48.63.120/Binding site: copper (His) #status predicted
K: keywords: copper; metalloprotein; oxidoreductase; zinc

F:57-146/Disulfide bonds: #status predicted
F:63,71,80,83/Binding site: zinc (His, His, Asp) #status predicted
F:143/Active site: Arg #status predicted

Query match	Score 659.3; DB 2; Length 135;
Best Local Similarity	83.0%; Pred. No. 1.9e-54;
Matches 127; Conservative	8; Mismatches 17; Indels 1; Gaps 1;

QY
1 MATKAVCVLKGDPVQGILNFEQESNGPVRVWGSIRGLTEGUGFHVFHFGDNTAGCTG 60
|||||
Db
1 MATKAVCVLKGDPVEATIHFEQK-GTGPVVVKGRITGLTEGLHFHVFHFGDNRRQGTG 59

DQ

60 AGPHFNP LSKRGKGGPKDERHVGDLGNVTADKGADVSIEDSVLSGDGHLIGRTLW 120
||||||| : ||||| : ||||| : ||||| :

Db

60 AGPHENP LSKRGGPKDERHVGDLGNVTAGNGVADVLIEDSVLSGDMSVGRTLW 119

```

QY      121 HKADDLGKGNEESTKTGNAGSRLACGVIGIA 153
      ||| |||||:|||||:|||||:
Db      120 HEKEDDLGKGNDSTKTGNAGSRLACGVIGIS 152

```


RESULT 8

DSBOCZ

superoxide dismutase (EC 1.15.1.1) (Cu-Zn) [validated] - bovine
N:Alternate names: thymus hypocholesterolemic factor
C:Species: Bos primigenius taurus (cattle)
C:Date: 24-Apr-1984 #sequence_revision 27-Feb-1997 #text_change 16-Jun-2000
C:Accession: I45883; S13674; A92163; A56902; A00513
R:Halliwell, R.A.; Imlay, K.C.; Lee, P.; Fong, N.M.; Gallegos, C.; Getzoff, E.D.; Tainer
Biochem. Biophys. Res. Commun. 181, 474-480, 1991
A:Title: Thermostabilization of recombinant human and bovine CuZn superoxide dismutases
A:Reference number: I45883; MUID:92068236; PMID:1958215
A:Accession: I45883
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-152 <HAL>
A:Cross-references: GB:M81129; NID:g162960; PIDN:AAA73164.1; PID:g162961
R:Gibbs, L.S.; Shaffer, J.B.
Nucleic Acids Res. 18, 7171, 1990
A:Title: Nucleotide sequence of bovine copper/zinc superoxide dismutase cDNA generated by
A:Reference number: S13674; MUID:91088332; PMID:2263495
A:Accession: S13674
A:Molecule type: mRNA
A:Residues: 2-152 <GIB>
A:Cross-references: EMBL:X54799
A:Experimental source: lung
R:Steinman, H.M.; Naik, V.R.; Abernethy, J.L.; Hill, R.L.
J. Biol. Chem. 249, 7326-7338, 1974
A:Title: Bovine erythrocyte superoxide dismutase. Complete amino acid sequence.
A:Reference number: A92163; MUID:75060425; PMID:4279916
A:Accession: A92163
A:Molecule type: protein
A:Residues: 2-152 <STE>
R:Abernethy, J.L.; Steinman, H.M.; Hill, R.L.
J. Biol. Chem. 249, 7339-7347, 1974
A:Title: Bovine erythrocyte superoxide dismutase. Subunit structure and sequence location
A:Reference number: A92164; MUID:75060426; PMID:4436313
A:Contents: annotation; disulfide bond
A>Note: the molecule is a dimer of identical chains
R:Richardson, J.S.; Thomas, K.A.; Rubin, B.H.; Richardson, D.C.
Proc. Natl. Acad. Sci. U.S.A. 72, 1349-1353, 1975
A:Title: Crystal structure of bovine Cu,Zn superoxide dismutase at 3 angstrom resolution
A:Reference number: A93796; MUID:75158289; PMID:1055410
A:Contents: annotation; X-ray crystallography; 3.0 angstroms
R:Tainer, J.A.; Getzoff, E.D.; Richardson, J.S.; Richardson, D.C.
submitted to the Brookhaven Protein Data Bank, March 1980
A:Reference number: A50536; PDB:2SOD
A:Contents: annotation; X-ray crystallography, 2.0 angstroms, residues 1-151
R:Malinowski, D.P.; Friedovich, I.
Biochemistry 18, 5909-5917, 1979
A:Title: Chemical modification of arginine at the active site of the bovine erythrocyte
A:Reference number: A90436; MUID:80088307; PMID:518876
A:Contents: annotation; active site
A>Note: chemical modification of Arg-142 inactivates the enzyme
R:Mondola, P.; Santillo, M.; Belfiore, A.; Camardella, L.; Santangelo, F.
Comp. Biochem. Physiol. B 104, 193-199, 1993
A:Title: The thymus hypocholesterolemic factor (TphF): a bovine thymic superoxide dismut
A:Reference number: A56902; MUID:93193420; PMID:8448990
A:Accession: A56902
A:Molecule type: protein
A:Residues: 5-38;70-130;136-152 <MON>
A>Note: sequence modified after extraction from NCBI backbone
C:Function:
A:Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxygen
C:Superfamily: superoxide dismutase (Cu-Zn)
C:Keywords: acetylated amino end; copper; homodimer; metalloprotein; oxidoreductase; zinc
F:2/Modified site: acetylated amino end (Ala) (in mature form) #status experimental
F:45,47,62,119/Binding site: copper (His) #status experimental
F:56-145/disulfide bonds: #status experimental
F:62,70,79,82/Binding site: zinc (His, Asp) #status experimental
F:142/Active site: Arg #status experimental

Query Match 79.4%; Score 652; DB 1; Length 152;

Best Local Similarity 81.88; Pred. No. 9.4e-54;

Matches 126; Conservative 10; Mismatches 16; Indels 2; Gaps 1;

QY 1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVHFEQDNTAGCTS 60

DB 1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVHFEQDNTAGCTS 58

QY 61 AGPHNPISLRKHGPKDERHVGDLGNVTADKGDVADVSIEDSVLSGDHCHIGRTLVV 120

DB 59 AGPHNPISLRKHGPKDERHVGDLGNVTADKGNVAIVDIDVPLTSLSGEYSIIIGRTMVV 118

QY 121 HEKADDLGKGGNESTKTGNAGSRLACGVIGTAQ 154

DB 119 HEKPDPLGRGGNESTKTGNAGSRLACGVIGIAK 152

RESULT 9

DSPGCGZ

superoxide dismutase (EC 1.15.1.1) (Cu-Zn) [validated] - pig

C:Species: Sus scrofa domestica (domestic pig)

C:Date: 04-Dec-1986 #sequence_revision 04-Dec-1986 #text_change 20-Apr-2000

C:Accession: A00514

R:Schinina, M.E.; Barra, D.; Simmaco, M.; Bossa, F.; Rottilio, G.

FEBS Lett. 186, 267-270, 1985

A:Title: Primary structure of porcine Cu,Zn superoxide dismutase.

A:Reference number: A00514; MUID:85231234; PMID:3891411

A:Accession: A00514

A:Molecule type: protein

A:Residues: 1-152 <SCH>

C:Function:

A:Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxyg

C:Superfamily: superoxide dismutase (Cu-Zn)

F:1/Modified site: acetylated amino end; copper; metalloprotein; oxidoreductase; zinc

F:45,47,62,119/Binding site: copper (His) #status predicted

F:56-145/disulfide bonds: #status predicted

F:62,70,79,82/Binding site: zinc (His, His, Asp) #status predicted

F:142/Active site: Arg #status predicted

Query Match 79.2%; Score 650.5; DB 1; Length 152;

Best Local Similarity 82.4%; Pred. No. 1.3e-53;

Matches 126; Conservative 7; Mismatches 19; Indels 1; Gaps 1;

QY 2 ATKAVCVLKGDGPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVHFEQDNTAGCTSA 61

DB 1 ATKAVCVLKGDGPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVHFEQDNTAGCTSA 59

QY 62 GFHFNPLSRKHGPKDERHVGDLGNVTADKGDVADVSIEDSVLSGDHCHIGRTLVVH 121

DB 60 GFHFNPESKKHGPKDERHVGDLGNVTAGDKGVATVYIEDSVIALSGDHSIIIGRTMVVH 119

QY 122 EKADDLGKGGNESTKTGNAGSRLACGVIGTAQ 154

DB 120 EKPDPLGRGGNESTKTGNAGSRLACGVIGITQ 152

RESULT 10

A24475

superoxide dismutase (EC 1.15.1.1) (Cu-Zn) - sheep

C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)

C:Date: 02-Jun-1988 #sequence_revision 02-Jun-1988 #text_change 05-Mar-1999

C:Accession: A24475

R:Schinina, M.E.; Barra, D.; Gentilomo, S.; Bossa, F.; Capo, C.; Rottilio, G.; Calabre

FEBS Lett. 207, 7-10, 1986

A:Title: Primary structure of a cationic Cu,Zn superoxide dismutase. The sheep enzyme

A:Reference number: A24475; MUID:87030907; PMID:3770194

A:Accession: A24475

A:Molecule type: protein

A:Residues: 1-151 <SCH>

C:Function:

A:Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxyg

C:Superfamily: superoxide dismutase (Cu-Zn)

C;Keywords: copper; metalloprotein; oxidoreductase; zinc
F;44,46,61,118/Binding site: copper (His) #status predicted
F;55-144/Disulfide bonds: #status predicted
F;61,69,78,81/Binding site: zinc (His, His, Asp) #status predicted
F;141/Active site: Arg #status predicted

Query Match 77.3%; Score 635; DB 2; Length 151;
Best Local Similarity 80.9%; Pred. No. 3.6e-52;
Matches 123; Conservative 8; Mismatches 19; Indels 2; Gaps 1;

QY 2 ATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVFHFGDNTAGCTSA 61
Db 1 ATKAVCVLKGDPVQGTITFEAK--GDKVVVGTSGITLGEDHGFHVFHFGDNTAGCTSA 58

QY 62 GPHFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSTEDSVISLSDHCHCIIGRTLVVH 121
Db 59 GPHFNPLSRKHGPKDEERHVGDLGNVKADKNGVAIVDPLISLSEYSIIIGRTMVVH 118

QY 122 EKADDLGKGGNEESTKTGNAGSLACGVIGIA 153
Db 119 EKPDLLGRGGNEESTKTGNAGSLACGVIGIA 150

RESULT 11
S65436
superoxide dismutase (EC 1.15.1.1) (Cu-Zn) - chicken
C;Species: Gallus gallus (chicken)
C;Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 20-Apr-2000
C;Accession: S65436
R;Schinina, M.E.; Carlini, P.; Politicelli, F.; Zappacosta, F.; Bossa, F.; Calabrese, L.
Eur. J. Biochem. 237, 433-439, 1996
A;Title: Amino acid sequence of chicken Cu, Zn-containing superoxide dismutase and ident
A;Reference number: S65436; MUID:96215439; PMID:8647082
A;Accession: S65436
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-153 <SCH>
C;Function:
A;Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxygen
C;Superfamily: superoxide dismutase (Cu-Zn)
C;Keywords: copper; metalloprotein; oxidoreductase; zinc
F;46,48,63,119/Binding site: copper (His) #status predicted
F;57-145/Disulfide bonds: #status predicted
F;63,71,80,83/Binding site: zinc (His, His, Asp) #status predicted
F;142/Active site: Arg #status predicted

Query Match 73.0%; Score 599; DB 2; Length 153;
Best Local Similarity 74.2%; Pred. No. 8.4e-49;
Matches 112; Conservative 19; Mismatches 18; Indels 2; Gaps 2;

QY 4 KAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVFHFGDNTAGCTSAGP 63
Db 4 KAVCVLKGDPVGVHVFQ--QSGPVKVTGKITGLSDGDHGFHVFHFGDNTAGCTSAGA 62

QY 64 HFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSTEDSVISLSDHCHCIIGRTLVVH 123
Db 63 HFNPEKGQHGPKDADRHVVDLGNVTA--KGGVAEVEIEDSVISLTPGHCIIIGRTMVVHAK 121

QY 124 ADDLKGKGGNEESTKTGNAGSLACGVIGIAQ 154
Db 122 SDDLGRGGNESKLTGNAGPRLACGVIGIAK 152

RESULT 12
S29782
superoxide dismutase (EC 1.15.1.1) (Cu-Zn) - loggerhead
C;Species: Caretta caretta (loggerhead)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 05-May-2000
C;Accession: S29782
R;Schinina, M.E.; Bossa, F.; Lania, A.; Capo, C.R.; Carlini, P.; Calabrese, L.
Eur. J. Biochem. 211, 843-849, 1993
A;Title: The primary structure of turtle Cu,Zn superoxide dismutase. Structural and func
A;Reference number: S29782; MUID:93170321; PMID:8436140

C;Keywords: copper; metalloprotein; oxidoreductase; zinc
F;44,46,61,118/Binding site: copper (His) #status predicted
F;55-144/Disulfide bonds: #status predicted
F;61,69,78,81/Binding site: zinc (His, His, Asp) #status predicted
F;141/Active site: Arg #status predicted

Query Match 77.3%; Score 635; DB 2; Length 151;
Best Local Similarity 80.9%; Pred. No. 3.6e-52;
Matches 123; Conservative 8; Mismatches 19; Indels 2; Gaps 1;

QY 2 ATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVFHFGDNTAGCTSA 61
Db 1 ATKAVCVLKGDPVQGTITFEAK--GDKVVVGTSGITLGEDHGFHVFHFGDNTAGCTSA 58

QY 62 GPHFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSTEDSVISLSDHCHCIIGRTLVVH 121
Db 59 GPHFNPLSRKHGPKDEERHVGDLGNVKADKNGVAIVDPLISLSEYSIIIGRTMVVH 118

QY 122 EKADDLGKGGNEESTKTGNAGSLACGVIGIA 153
Db 119 EKPDLLGRGGNEESTKTGNAGSLACGVIGIA 150

RESULT 11
S65436
superoxide dismutase (EC 1.15.1.1) (Cu-Zn) - chicken
C;Species: Gallus gallus (chicken)
C;Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 20-Apr-2000
C;Accession: S65436
R;Schinina, M.E.; Carlini, P.; Politicelli, F.; Zappacosta, F.; Bossa, F.; Calabrese, L.
Eur. J. Biochem. 237, 433-439, 1996
A;Title: Amino acid sequence of chicken Cu, Zn-containing superoxide dismutase and ident
A;Reference number: S65436; MUID:96215439; PMID:8647082
A;Accession: S65436
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-153 <SCH>
C;Function:
A;Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxygen
C;Superfamily: superoxide dismutase (Cu-Zn)
C;Keywords: copper; metalloprotein; oxidoreductase; zinc
F;46,48,63,119/Binding site: copper (His) #status predicted
F;57-145/Disulfide bonds: #status predicted
F;63,71,80,83/Binding site: zinc (His, His, Asp) #status predicted
F;142/Active site: Arg #status predicted

Query Match 73.0%; Score 599; DB 2; Length 153;
Best Local Similarity 74.2%; Pred. No. 8.4e-49;
Matches 112; Conservative 19; Mismatches 18; Indels 2; Gaps 2;

QY 4 KAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVFHFGDNTAGCTSAGP 63
Db 4 KAVCVLKGDPVGVHVFQ--QSGPVKVTGKITGLSDGDHGFHVFHFGDNTAGCTSAGA 62

QY 64 HFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSTEDSVISLSDHCHCIIGRTLVVH 123
Db 63 HFNPEKGQHGPKDADRHVVDLGNVTA--KGGVAEVEIEDSVISLTPGHCIIIGRTMVVHAK 121

QY 124 ADDLKGKGGNEESTKTGNAGSLACGVIGIAQ 154
Db 122 SDDLGRGGNESKLTGNAGPRLACGVIGIAK 152

RESULT 12
S29782
superoxide dismutase (EC 1.15.1.1) (Cu-Zn) - loggerhead
C;Species: Caretta caretta (loggerhead)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 05-May-2000
C;Accession: S29782
R;Schinina, M.E.; Bossa, F.; Lania, A.; Capo, C.R.; Carlini, P.; Calabrese, L.
Eur. J. Biochem. 211, 843-849, 1993
A;Title: The primary structure of turtle Cu,Zn superoxide dismutase. Structural and func
A;Reference number: S29782; MUID:93170321; PMID:8436140

A;Accession: S29782
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-166 <SCH>
C;Function:
A;Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxyg
C;Superfamily: superoxide dismutase (Cu-Zn)
C;Keywords: copper; metalloprotein; oxidoreductase; zinc
F;58,60,75,132/Binding site: copper (His) #status predicted
F;69-158/Disulfide bonds: #status predicted
F;75,83,92,95/Binding site: zinc (His, His, Asp) #status predicted
F;155/Active site: Arg #status predicted

Query Match 71.8%; Score 589.5; DB 2; Length 166;
Best Local Similarity 69.9%; Pred. No. 7.2e-48;
Matches 114; Conservative 16; Mismatches 20; Indels 13; Gaps 2;

QY 4 KAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVFHFGDNTAGCTSA 61
Db 4 KAVCVLKGDPVKEPVKVPKVPKVIYFEQ--QNGPVTLSGSIITGLTEGKHGFHVFHFGD 62

QY 52 GDNATAGCTSAGPHENPLSRKHGPKDEERHVGDLGNVTADKGVADVSTEDSVISLSDH 111
Db 63 GDNATAGCTSAGAHENPPGKNHGGPDNERHVGDLGNVTADKGVADVSTEDSVISLSDH 112

QY 112 CIIGRTLVVHEKADDLGKGGNEESTKTGNAGSLACGVIGIAQ 154
Db 123 SIIGRTMVVHEKEDDLGKGGNDESLKTGNAGSLACGVIGIAK 165

RESULT 13
S04623
superoxide dismutase (EC 1.15.1.1) (Cu-Zn) - blue shark
C;Species: Prionace glauca (blue shark)
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 20-Apr-2000
C;Accession: S04623; S45642
R;Calabrese, L.; Politicelli, F.; O'Neill, P.; Galtieri, A.; Barra, D.; Schinina, E.;
FEBS Lett. 250, 49-52, 1989
A;Title: Substitution of arginine for lysine 134 alters electrostatic parameters of t
A;Reference number: S04623; MUID:89250032; PMID:2500367
A;Accession: S04623
A;Molecule type: protein
A;Residues: 1-152 <CAL>
C;Function:
A;Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxyg
C;Superfamily: superoxide dismutase (Cu-Zn)
C;Keywords: copper; metalloprotein; oxidoreductase; zinc
F;44,46,61,118/Binding site: copper (His) #status predicted
F;55-144/Disulfide bonds: #status predicted
F;61,69,78,81/Binding site: zinc (His, His, Asp) #status predicted
F;141/Active site: Arg #status predicted

Query Match 68.0%; Score 558.5; DB 2; Length 152;
Best Local Similarity 70.9%; Pred. No. 5.1e-45;
Matches 107; Conservative 15; Mismatches 28; Indels 1; Gaps 1;

QY 4 KAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVFHFGDNTAGCTSAGP 63
Db 2 KAVCVLKGDTGVTGTVLFEQ--ADGPVTLKGSITGLTPGKHGFHVFHFGDNTAGCTSAGP 60

QY 64 HFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSTEDSVISLSDHCHCIIGRTLVVH 123
Db 61 HYNPFKSNHGGPDDEERHVGDLGNVEANGVAFEIKRQLHLSGERSIIGRTLVVH 120

QY 124 ADDLKGKGGNEESTKTGNAGSLACGVIGIAQ 154
Db 121 EDDLKGKGGDEESLRTGNAGSLACGVIGIAK 151

RESULT 14
DSWFCZ
superoxide dismutase (EC 1.15.1.1) (Cu-Zn) - swordfish
C;Species: Xiphias gladius (swordfish)

C:Date: 13-Aug-1986 #sequence_revision 13-Aug-1986 #text_change 05-Mar-1999

C:Accession: A00516

R:Rocha, H.A.; Bannister, W.H.; Bannister, J.V.

Eur. J. Biochem. 145, 477-484, 1984

A:Title: The amino-acid sequence of copper/zinc superoxide dismutase from swordfish liver

A:Reference number: A00516; MUID:85076642; PMID:6510412

A:Accession: A00516

A:Molecule type: protein

A:Residues: 1-151 <ROC>

C:Function:

A:Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxygen

C:Superfamily: superoxide dismutase (Cu-Zn)

C:Keywords: copper; metalloprotein; oxidoreductase; zinc

F:46,48,63,119/Binding site: copper (His) #status predicted

F:57-145/Disulfide bonds: #status predicted

F:63,71,80,83/Binding site: zinc (His, His, Asp) #status predicted

F:142/Active site: Arg #status predicted

Query Match

Best Local Similarity 65.3%; Score 536.5; DB 1; Length 151;

Matches 103; Conservative 14; Mismatches 30; Indels 1; Gaps 1;

QY 4 KAVCVLKGDPVQGIINFQKESNGPVKVGSIKGLTEGLHGFHVHFGDNTAGCTSAGP 63

Db 3 KAVCVLKGDPVQGIINFQKESNGPVKVGSIKGLTEGLHGFHVHFGDNTAGCTSAGP 63

QY 64 HFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSGDHCHCIIGRTLVVHEK 123

Db 63 HFNPKKHAGPKDEERHVGDLGNVTADANGVAKIDITDK-ISTLPGYSIIIGRTMVIHEK 121

QY 124 ADDLKGKGNBESLTKGNAGSRLACGVIG 151

Db 122 ADDLGRGNGEESLTKGNAGSRLACGVIG 149

RESULT 15

S09568

Superoxide dismutase (EC 1.15.1.1) (Cu-Zn) B - African clawed frog

C:Species: Xenopus laevis (African clawed frog)

C:Date: 13-Jan-1995 #sequence_revision 11-Apr-1997 #text_change 11-Jun-1999

C:Accession: S09568; S05022; S59617

R:Carri, M.T.; Battistoni, A.; Mariottini, P.; Rotilio, G.

Nucleic Acids Res. 18, 1641, 1990

A:Title: Xenopus laevis Cu,Zn superoxide dismutase B cDNA sequence.

A:Reference number: S09568; MUID:90221905; PMID:2326205

A:Accession: S09568

A:Molecule type: mRNA

A:Residues: 1-151 <CAR>

A:Cross-references: EMBL:X51518; NID:g65258; PIDN:CAA35890.1; PID:g65259

R:Schinina, M.E.; Barra, D.; Bossa, F.; Calabrese, L.; Montesano, L.; Carri, M.T.; Mariottini, P.

Arch. Biochem. Biophys. 272, 507-515, 1989

A:Title: Primary structure from amino acid and cDNA sequences of two Cu,Zn superoxide dismutases from

A:Reference number: S05021; MUID:89321563; PMID:2751312

A:Accession: S05022

A:Molecule type: protein

A:Residues: 2-60, 'P', 62-151 <SCH>

A:Note: residues 78-151 were also determined by cDNA sequencing

A:Accession: S59617

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 78-151 <SC2>

C:Function:

A:Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxygen

C:Superfamily: superoxide dismutase (Cu-Zn)

C:Keywords: copper; metalloprotein; oxidoreductase; zinc

F:45,47,62,118/Binding site: copper (His) #status predicted

F:56-144/Disulfide bonds: #status predicted

F:62,70,79,82/Binding site: zinc (His, His, Asp) #status predicted

F:141/Active site: Arg #status predicted

Query Match

Best Local Similarity 64.9%; Score 533; DB 1; Length 151;

Matches 102; Conservative 16; Mismatches 28; Indels 2; Gaps 2;

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:03:18 ; Search time 2.35564 Seconds
(without alignments)
2711.515 Million cell updates/sec

Title: US-09-904-987-6
Perfect score: 821
Sequence: 1 MATRAVCVLKGDGPVQGIIN.....STKTGNAGSLACGVIGIAQ 154

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	816	99.4	153	1 SODC_HUMAN	P00441 homo sapien
2	684.5	83.4	152	1 SODC_CAVPO	P33431 cavia porce
3	675	82.2	153	1 SODC_MOUSE	P08228 mus musculu
4	673	82.0	153	1 SODC_RAT	P07632 rattus norv
5	662	80.6	153	1 SODC_HORSE	P00443 equus cabal
6	654.5	79.7	152	1 SODC_RABIT	P09212 oryctolagus
7	651	79.3	151	1 SODC_CEREL	O46412 cervus elap
8	650.5	79.2	152	1 SODC_PIG	P04178 sus scrofa
9	647	78.8	151	1 SODC_BOVIN	P00442 bos taurus
10	635	77.3	151	1 SODC_SHEEP	P09670 ovis aries
11	593	72.2	153	1 SODC_CHICK	P80566 gallus gall
12	589.5	71.8	166	1 SODC_CARGR	P80174 caretta car
13	579	70.5	154	1 SODC_BRARE	O73872 brachydanio
14	558.5	68.0	152	1 SODC_PRIGL	P14118 prionace gl
15	541	65.9	150	1 SODD_XENLA	P15107 xenopus lae
16	536.5	65.3	151	1 SODC_XIPGL	P03946 xiphias gla
17	521	63.5	150	1 SODC_XENLA	P13926 xenopus lae
18	507	61.8	152	1 SODC_CERCA	P28755 ceratitlis c
19	500	60.9	152	1 SODC_CHYAM	Q07182 chymomyza a
20	497	60.5	153	1 SODC_SCHMA	Q11137 schistosoma
21	495	60.3	152	1 SODC_DROWI	P141973 drosophila
22	493.5	60.1	159	1 SODC_HAECO	Q27666 haemonchus
23	488	59.4	152	1 SODC_DROME	P00444 drosophila
24	483.5	58.9	153	1 SODC_DEBHA	O42724 debaryomyce
25	482.5	58.8	154	1 SODC_PINSY	P24669 pinus sylve
26	482	58.7	152	1 SODC_DROVI	P10791 drosophila
27	480	58.5	150	1 SODC_MAIZE	P11428 zea mays (m
28	477.5	58.2	151	1 SOD2_ORYSA	P28757 oryza sativ
29	474	57.7	158	1 SODC_BRUPA	P41962 brugia paha
30	473.5	57.7	220	1 SODP_SOLCS	O04997 solidago ca
31	472.5	57.6	151	1 SODC_HALRO	P81926 halocynthia
32	472.5	57.6	153	1 SODC_ASPTU	Q9y849 aspergillus
33	470.5	57.3	151	1 SOD5_MAIZE	P23346 zea mays (m

RESULT 1				
SODC_HUMAN				
ID	SODC_HUMAN	STANDARD;	PRT;	153 AA.
AC	P00441; Q16711; Q16669; Q16838; Q16839; Q16840;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Superoxide dismutase [Cu-Zn] (EC 1.15.1.1).			
GN	SOD1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=85257452; PubMed=3160582;			
RA	Levanon D., Lieman-Hurwitz J., Dafni N., Wigderson M., Sherman L.,			
RA	Bernstein Y., Laver-Rudich Z., Danciger E., Stein O., Groner Y.;			
RT	"Architecture and anatomy of the chromosomal locus in human			
RT	chromosome 21 encoding the Cu/Zn superoxide dismutase.";			
RL	EMBO J. 4:77-84(1985).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=85215596; PubMed=3889846;			
RA	Hallewell R.A., Masiazk F.R., Najarian R.C., Puma J.P., Quiroga M.R.,			
RA	Randolph A., Sanchez-Pescador R., Scandella C.J., Smith B.,			
RA	Steimer K.S., Mullenbach G.T.;			
RT	"Human Cu/Zn superoxide dismutase cDNA: isolation of clones			
RT	synthesising high levels of active or inactive enzyme from an			
RT	expression library.";			
RL	Nucleic Acids Res. 13:2017-2034(1985).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=83299994; PubMed=6577438;			
RA	Sherman L., Dafni N., Lieman-Hurwitz J., Groner Y.;			
RT	"Nucleotide sequence and expression of human chromosome 21-encoded			
RT	superoxide dismutase mRNA.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 80:5465-5469(1983).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=89174523; PubMed=2853161;			
RA	Kajihara J., Enomoto M., Nishijima K., Yabuuchi M., Katoh K.;			
RT	"Comparison of properties between human recombinant and placental			
RT	copper-zinc SOD.";			
RL	J. Biochem. 104:851-854(1988).			
RN	[5]			
RP	SEQUENCE.			
RX	MEDLINE=81067132; PubMed=7002610;			
RA	Barra D., Martini F., Bannister J.V., Schinina M.E., Rotillo G.,			
RA	Bannister W.H., Bossa F.;			
RT	"The complete amino acid sequence of human Cu/Zn superoxide			
RT	dismutase.";			
RL	FEBS Lett. 120:53-56(1980).			
RN	[6]			
RP	SEQUENCE.			
RX	MEDLINE=80221052; PubMed=6770891;			

P23345 zea mays (m
O65198 medicago sa
P54407 drosophila
O22668 panax glnse
O07449 onchocerca
O78310 arabidopsis
P07505 spinacia ol
P11964 pisum sativ
P93407 oryza sativ
P28756 oryza sativ
P24706 onchocerca
Q98q15 ananas como

- RA Jabusch J.R., Farb D.L., Kerschensteiner D.A., Deutsch H.F.;
 RT "some sulfhydryl properties and primary structure of human
 RT erythrocyte superoxide dismutase.";
 RL Biochemistry 19:2310-2316(1980).
 RN [7]
- RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
 RX MEDLINE=92335247; PubMed=1463506;
 RA Parge H.E., Hallewell R.A., Tainer J.A.;
 RT "Atomic structures of wild-type and thermostable mutant recombinant
 RT human Cu,Zn superoxide dismutase.";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:6109-6113(1992).
 RN [8]
- RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF VARIANT GLY-37.
 RX MEDLINE=98200458; PubMed=9541385;
 RA Hart P.J., Liu H., Pellegrini M., Nersissian A.M., Gralla E.B.,
 RA Valentine J.S., Eisenberg D.;
 RT "Subunit asymmetry in the three-dimensional structure of a human
 RT Cu,ZnSOD mutant found in familial amyotrophic lateral sclerosis.";
 RL Protein Sci. 7:545-555(1998).
 RN [9]
- RP X-RAY CRYSTALLOGRAPHY (1.02 ANGSTROMS).
 RX MEDLINE=99262851; PubMed=10329151;
 RA Ferraroni M., Rypniewski W., Wilson K.S., Viezzoli M.S., Banci L.,
 RA Bertini I., Mangani S.;
 RT "The crystal structure of the monomeric human SOD mutant
 RT F50E/G51E/E133Q at atomic resolution. The enzyme mechanism
 RT revisited.";
 RL J. Mol. Biol. 288:413-426(1999).
 RN [10]
- RP STRUCTURE BY NMR.
 RX MEDLINE=98384143; PubMed=9718300;
 RA Banci L., Benedetto M., Bertini I., del Conte R., Piccioli M.,
 RA Viezzoli M.S.;
 RT "Solution structure of reduced monomeric Q133M2 copper, zinc
 RT superoxide dismutase (SOD). Why is SOD a dimeric enzyme?;
 RL Biochemistry 37:11780-11791(1998).
 RN [11]
- RP REVIEW ON VARIANTS.
 RX MEDLINE=96154858; PubMed=8592323;
 RA de Bellecoche J., Orrell R., King A.;
 RT "Familial amyotrophic lateral sclerosis/motor neurone disease (FALS):
 RT a review of current developments.";
 RL J. Med. Genet. 32:841-847(1995).
 RN [12]
- RP VARIANTS FALS.
 RX MEDLINE=93188958; PubMed=8446170;
 RA Rosen D.R., Siddique T., Patterson D., Figlewicz D.A., Sapp P.,
 RA Hentati A., Donaldson D., Goto J., O'Regan J.P., Deng H.-X.,
 RA Rahmani Z., Krizus A., McKenna-Yasek D., Cayabyab A., Gaston S.M.,
 RA Berger R., Tanzi R.E., Halperin J.J., Herzfeldt B., van den Bergh R.,
 RA Hung W.-Y., Bird T.E., Deng G., Mulder D.W., Smyth C., Laing N.G.,
 RA Soriano E., Pericak-Vance M.A., Haines J., Rouleau G.A., Gusella J.S.,
 RA Horvitz H.R., Brown R.H. Jr.;
 RT "Mutations in Cu,Zn superoxide dismutase gene are associated with
 RT familial amyotrophic lateral sclerosis.";
 RL Nature 362:59-62(1993).
 RN [13]
- RP ERRATUM.
 RX MEDLINE=93323981; PubMed=8332197;
 RA Rosen D.R.;
 RL Nature 364:362-362(1993).
 RN [14]
- RP VARIANTS FALS.
 RX MEDLINE=93355289; PubMed=8351519;
 RA Deng H.-X., Hentati A., Tainer J.A., Iqbal Z., Cayabyab A.,
 RA Hung W.-Y., Getzoff E.D., Hu P., Herzfeldt B., Roos R.P., Warner C.,
 RA Deng G., Soriano E., Smyth C., Parge H.E., Ahmed A., Roses A.D.,
 RA Hallewell R.A., Pericak-Vance M.A., Siddique T.;
 RT "Amyotrophic lateral sclerosis and structural defects in Cu,Zn
 RT superoxide dismutase.";
 RL Science 261:1047-1051(1993).
 RN [15]
- RP VARIANT FALS THR-4.
 RX MEDLINE=94235014; PubMed=8179602;
 RA Nakano R., Sato S., Inuzuka T., Sakimura K., Mishina M., Takahashi H.,
 RA Ikuta F., Honma Y., Fujii J., Taniguchi N., Tsuji S.;
 RT "A novel mutation in Cu,Zn superoxide dismutase gene in Japanese
 RT familial amyotrophic lateral sclerosis.";
 RL Biochem. Biophys. Res. Commun. 200:695-703(1994).
 RN [16]
- RP VARIANT FALS GLU-7.
 RX MEDLINE=95071364; PubMed=7980516;
 RA Hirano M., Fujii J., Nagai Y., Sonobe M., Okamoto K., Araki H.,
 RA Taniguchi N., Ueno S.;
 RT "A new variant Cu,Zn superoxide dismutase (Val7-->Glu) deduced from
 RT lymphocyte mRNA sequences from Japanese patients with familial
 RT amyotrophic lateral sclerosis.";
 RL Biochem. Biophys. Res. Commun. 204:572-577(1994).
 RN [17]
- RP VARIANT FALS LYS-21.
 RX MEDLINE=94348517; PubMed=8069312;
 RA Jones C.T., Swinger R.J., Brock D.J.H.;
 RT "Identification of a novel SOD1 mutation in an apparently sporadic
 RT amyotrophic lateral sclerosis patient and the detection of Ile113Thr
 RT in three others.";
 RL Hum. Mol. Genet. 3:649-650(1994).
 RN [18]
- RP VARIANTS FALS ASP-93 AND THR-112.
 RX MEDLINE=95038761; PubMed=7951252;
 RA Esteban J., Rosen D.R., Bowling A.C., Sapp P., McKenna-Yasek D.,
 RA O'Regan J.P., Beal M.F., Horvitz H.R., Brown R.H. Jr.;
 RT "Identification of two novel mutations and a new polymorphism in the
 RT gene for Cu,Zn superoxide dismutase in patients with amyotrophic
 RT lateral sclerosis.";
 RL Hum. Mol. Genet. 3:997-998(1994).
 RN [19]
- RP VARIANT FALS GLY-115.
 RX MEDLINE=95187174; PubMed=7881433;
 RA Kostzewa M., Burck-Lehmann U., Mueller U.;
 RT "Autosomal dominant amyotrophic lateral sclerosis: a novel mutation
 RT in the Cu,Zn superoxide dismutase-1 gene.";
 RL Hum. Mol. Genet. 3:2261-2262(1994).
 RN [20]
- RP VARIANT FALS ARG-46.
 RX MEDLINE=95138736; PubMed=7836951;
 RA Aoki M., Ogasawara M., Matsubara Y., Narisawa K., Nakamura S.,
 RA Itoyama Y., Abe K.;
 RT "Familial amyotrophic lateral sclerosis (ALS) in Japan associated
 RT with H46R mutation in Cu,Zn superoxide dismutase gene: a possible new
 RT subtype of familial ALS.";
 RL J. Neurol. Sci. 126:77-83(1994).
 RN [21]
- RP VARIANT FALS THR-113.
 RX MEDLINE=95089537; PubMed=7997024;
 RA Suthers G., Laing N., Wilton S., Dorosz S., Waddy H.;
 RT "Sporadic motoneuron disease due to familial SOD1 mutation with low
 RT penetrance.";
 RL Lancet 344:1773-1773(1994).
 RN [22]
- RP VARIANT FALS ASN-101.
 RX MEDLINE=95174778; PubMed=7870076;
 RA Jones C.T., Shaw P.J., Chari G., Brock D.J.;
 RT "Identification of a novel exon 4 SOD1 mutation in a sporadic
 RT amyotrophic lateral sclerosis patient.";
 RL Mol. Cell. Probes 8:329-330(1994).
 RN [23]
- RP VARIANTS FALS.
 RX MEDLINE=95193785; PubMed=7887412;
 RA Pramatarova A., Figlewicz D.A., Krizus A., Han F.Y.,
 RA Ceballos-Picot I., Nicole A., Dib M., Meininger V., Brown R.H.,
 RA Rouleau G.A.;
 RT "Identification of new mutations in the Cu,Zn superoxide dismutase
 RT gene of patients with familial amyotrophic lateral sclerosis.";
 RL Am. J. Hum. Genet. 56:592-596(1995).
 RN [24]
- RP VARIANT FALS ILB-148.

Query Match 99.4%; Score 816; DB 1; Length 153;
 Best Local Similarity 100.0%; Pred. No. 1e-68;
 Matches 133; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ATKAVCVLKGDGPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVFHFGDNTAGCTSA 61
 |||||
 DB 1 ATKAVCVLKGDGPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVFHFGDNTAGCTSA 60
 |||||

QY 62 GPHEFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSIEDSVISLGDHCHIIIGRTLTVVH 121
 |||||
 DB 61 GPHEFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSIEDSVISLGDHCHIIIGRTLTVVH 120
 |||||

QY 122 EKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
 |||||
 DB 121 EKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 153
 |||||

RESULT 2
 SODC_CAVPO STANDARD; PRT; 152 AA.
 AC P33431;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Superoxide dismutase [Cu-Zn] (EC 1.15.1.1).
 GN SOD1.
 OS Cavia porcellus (Guinea pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
 OX NCBI_TaxID=10141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Hartley; TISSUE=Lung;
 RA Yuan H.T., Bingle C.D., Kelly F.J.;
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE.
 RC TISSUE=Liver;
 RX MEDLINE=94059414; PubMed=8240718;
 RA Wolf B., Reinecke K., Aumann K.-D., Brigelius-Flohe R., Flohe L.;
 RT "Taxonomical classification of the guinea pig based on its Cu/Zn
 superoxide dismutase sequence."
 RL Biol. Chem. Hoppe-Seyler 374:641-649(1993).
 CC -|- FUNCTION: Destroys radicals which are normally produced within the
 CC cells and which are toxic to biological systems.
 CC -|- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
 CC -|- COFACTOR: Copper and zinc.
 CC -|- SUBUNIT: HOMODIMER.
 CC -|- SIMILARITY: BELONGS TO THE CU-ZN SUPEROXIDE DISMUTASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U39844; AAC52720.1; -
 CC DR HSP: P00441; 4SOD.
 CC DR InterPro: IPR001424; SOD_CU_ZN.
 CC DR PRINTS: PF00080; sodcu.1.
 CC DR PRODOM: PD000469; SOD_CU_ZN.1.
 CC DR PROSITE: PS00087; SOD_CU_ZN.1; 1.
 CC DR PROSITE: PS00332; SOD_CU_ZN.2; 1.
 CC Oxidoreductase; Copper; Zinc; Acetylation.
 KW INIT_MET 0 0
 FT MOD_RES 1 1 ACETYLATION.
 FT METAL 45 45 COPPER (BY SIMILARITY).
 FT

FT METAL 47 47 COPPER (BY SIMILARITY).
 FT METAL 62 62 COPPER AND ZINC (BY SIMILARITY).
 FT METAL 70 70 ZINC (BY SIMILARITY).
 FT METAL 79 79 ZINC (BY SIMILARITY).
 FT METAL 82 82 ZINC (BY SIMILARITY).
 FT METAL 119 119 COPPER (BY SIMILARITY).
 FT DISULFID 56 145 BY SIMILARITY.
 FT CONFLICT 102 103 LI -> IL (IN REF. 2).
 SQ SEQUENCE 152 AA; 15538 MW; 5891CE16E2ACBBF3 CRC64;

Query Match 83.4%; Score 684.5; DB 1; Length 152;
 Best Local Similarity 86.3%; Pred. No. 1.5e-56;
 Matches 132; Conservative 8; Mismatches 12; Indels 1; Gaps 1;

QY 2 ATKAVCVLKGDGPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVFHFGDNTAGCTSA 61
 |||||
 DB 1 ATKAVCVLKGDGPVQGIINFEQK-ANGPVVYKGRITGLVEGKHGPHVHFGDNTAGCTSA 59
 |||||

QY 62 GPHEFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSIEDSVISLGDHCHIIIGRTLTVVH 121
 |||||
 DB 60 GPHEFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSIEDSVISLGSANSTIGRTVMVH 119
 |||||

QY 122 EKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
 |||||
 DB 120 EXPDDLKGGNEESTKTGNAGSRLACGVIGIAQ 152
 |||||

RESULT 3
 SODC_MOUSE STANDARD; PRT; 153 AA.
 AC P08228;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Superoxide dismutase [Cu-Zn] (EC 1.15.1.1).
 GN SOD1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SWR/J; TISSUE=Liver;
 RX MEDLINE=88203220; PubMed=3362683;
 RA Bewley G.C.;
 RT "cDNA and deduced amino acid sequence of murine Cu-Zn superoxide
 RT dismutase."
 RL Nucleic Acids Res. 16:2728-2728(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91216458; PubMed=2022332;
 RA Benedetto M.T., Anzai Y., Gordon J.W.;
 RT "Isolation and analysis of the mouse genomic sequence encoding
 RT Cu(2+)-Zn2+ superoxide dismutase."
 RL Gene 99:191-195(1991).
 RN [3]
 RP SEQUENCE OF 3-22.
 RX MEDLINE=90361747; PubMed=2391363;
 RA Pluthero F.G., Shreeve M., Eskinazi D., van der Gaag H., Huang K.S.,
 RA Hulmes J.D., Blum M., Axelrad A.A.;
 RT "Purification of an inhibitor of erythroid progenitor cell cycling
 RT and antagonist to interleukin 3 from mouse marrow cell supernatants
 RT and its identification as cytosolic superoxide dismutase."
 RL J. Cell Biol. 111:1217-1223(1990).
 CC -|- FUNCTION: Destroys radicals which are normally produced within the
 CC cells and which are toxic to biological systems.
 CC -|- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
 CC -|- COFACTOR: Copper and zinc.
 CC -|- SUBUNIT: HOMODIMER.
 CC -|- SIMILARITY: BELONGS TO THE CU-ZN SUPEROXIDE DISMUTASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: X06683; CAA29880.1; -
 DR EMBL: M60798; AAA40121.1; -
 DR EMBL: M60794; AAA40121.1; JOINED.
 DR EMBL: M60795; AAA40121.1; JOINED.
 DR EMBL: M60796; AAA40121.1; JOINED.
 DR EMBL: M60797; AAA40121.1; JOINED.
 DR EMBL: M35725; AAA37518.1; -
 DR PIR: A37101; A37101.
 DR PIR: JQ0915; JQ0915.
 DR HSSP: P00442; ICBJ.
 DR SWISS-2DPAGE; P08228; MOUSE.
 DR MGD; MGI:98351; Sod1.
 DR InterPro; IPR001424; SOD_CU_ZN.
 DR Pfam; PF00080; sodcu; 1.
 DR PRINTS; PR00068; CUZNDISWASE.
 DR ProDom; PD000469; SOD_CU_ZN; 1.
 DR PROSITE; PS00087; SOD_CU_ZN_1; 1.
 DR PROSITE; PS00332; SOD_CU_ZN_2; 1.
 KW Oxidoreductase; Copper; Zinc.
 FT INIT.MET 0
 FT METAL 46
 FT METAL 46
 FT METAL 48
 FT METAL 63
 FT METAL 71
 FT METAL 80
 FT METAL 83
 FT METAL 120
 FT METAL 120
 FT DISULFID 57
 FT CONFLICT 101 101
 SQ SEQUENCE 153 AA; 15811 MW; 6305F51B44DEBA8D CRC64;

Query Match 82.2%; Score 675; DB 1; Length 153;
 Best Local Similarity 83.7%; Pred. No. 1.le-55;
 Matches 128; Conservative 8; Mismatches 17; Indels 0; Gaps 0;
 QY 2 ATKAVCVLKGDPVQGIIFEQKESNGPVKVGSIKGLTEGLHGFHVEFGDNTAGCTSA 61
 Db 1 AMKAVCVLKGDPVQGTIFHEQKASGPVVLGSGITGLTEGQGRHVHVGNDTGCTSA 60
 QY 62 GPHFNPLSRKHGPKDEERHVGDLGNVTADKGVADYSDVSLSGDHCIIIGRTLTVH 121
 Db 61 GPHFNPHSKKHGPKDEERHVGDLGNVTAGKGVANVSDRVISLSGHSIIIGRTMVVH 120
 QY 122 EKADDLKKGNEESTKTGNAGSLACGVIGIAQ 154
 Db 121 EKQDDLKKGNEESTKTGNAGSLACGVIGIAQ 153

RESULT 4
 ID SODC_RAT STANDARD; PRT; 153 AA.
 AC P07632;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-APR-1988 (Rel. 07, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Superoxide dismutase [Cu-Zn] (EC 1.15.1.1).
 GN SOD1
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OC NCBI_TaxId=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Sprague-Dawley; TISSUE=Liver;
 RX MEDLINE=87316943; PubMed=3628012;
 RA HO Y.-S., Crapo J.D.;

RT "cDNA and deduced amino acid sequence of rat copper-zinc-containing
 RT superoxide dismutase.";
 RL Nucleic Acids Res. 15:6746-6746(1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RX MEDLINE=89198078; PubMed=2703531;
 RA Hass M.A., Iqbal J., Clerch L.B., Frank L., Massaro D.;
 RT "Rat lung Cu,Zn superoxide dismutase. Isolation and sequence of a
 RL full-length cDNA and studies of enzyme induction.";
 RN J. Clin. Invest. 83:1241-1246(1989).
 RP [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92360048; PubMed=1379810;
 RA Hsu J.L., Visner G.A., Burr I.A., Nick H.S.;
 RT "Rat copper/zinc superoxide dismutase gene: isolation,
 RL characterization, and species comparison.";
 RN Biochem. Biophys. Res. Commun. 186:936-943(1992).
 RP [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Sprague-Dawley; TISSUE=Liver;
 RX MEDLINE=94040822; PubMed=8224914;
 RA Kim Y.H., Yoo H.Y., Jung G., Kim J.Y., Rho H.M.;
 RT "Isolation and analysis of the rat genomic sequence encoding Cu/Zn
 RL superoxide dismutase.";
 RN Gene 133:267-271(1993).
 RP [5]
 RP SEQUENCE OF 3-153 FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=87246675; PubMed=3595611;
 RA Delabar J.-M., Nicole A., D'Aurion L., Jacob Y., Meunier-Rotival M.,
 RN Galibert F., Sinet P.-M., Jerome H.;
 RT "Cloning and sequencing of a rat CuZn superoxide dismutase cDNA.
 RL Correlation between CuZn superoxide dismutase mRNA level and enzyme
 RN activity in rat and mouse tissues.";
 RP Eur. J. Biochem. 166:181-187(1987).
 RP [6]
 RP SEQUENCE.
 RX MEDLINE=87076037; PubMed=3790250;
 RA Steffens G.J., Michelson A.M., Puget K., Flohe L.;
 RT "The amino-acid sequence of rat Cu-Zn superoxide dismutase.";
 RL Biol. Chem. Hoppe-Seyler 367:1017-1024(1986).
 RN [7]
 RP SEQUENCE OF 1-20 AND 92-102.
 RX MEDLINE=96434889; PubMed=8837775;
 RA Wang X., Culotta V.C., Klee C.B.;
 RT "Superoxide dismutase protects calcein from inactivation.";
 RN Nature 383:434-437(1996).
 CC -!- FUNCTION: Destroys radicals which are normally produced within the
 CC cells and which are toxic to biological systems.
 CC -!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) -> O(2) + H(2)O(2).
 CC -!- COFACTOR: Copper and zinc.
 CC -!- SUBUNIT: HOMODIMER.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- SIMILARITY: BELONGS TO THE CU-ZN SUPEROXIDE DISMUTASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: Y00404; CAA68465.1; -
 DR EMBL: M25157; AAA42160.1; ALT_INIT.
 DR EMBL: X05634; CAA29121.1; -
 DR EMBL: M21917; CAA79925.1; ALT_INIT.
 DR EMBL: M21918; CAA79925.1; JOINED.
 DR EMBL: M21919; CAA79925.1; JOINED.
 DR EMBL: M21920; CAA79925.1; JOINED.
 DR PIR: S00036; S00036.
 DR PIR: JC1192; JC1192.


```

DR PIR; JC2015; JC2015.
DR HSP; P00442; ICBJ.
DR InterPro: IPR001424; SOD_CU_ZN.
DR Pfam; PF00080; sodcu; 1.
DR PRINTS; PR00068; CUZNDISMTASE.
DR ProDom; PD000469; SOD_CU_ZN; 1.
DR PROSITE; PS00087; SOD_CU_ZN_1; 1.
DR PROSITE; PS00332; SOD_CU_ZN_2; 1.
KW Oxidoreductase; Copper; Zinc.
FT INIT_MET 0
FT METAL 46 46 COPPER (BY SIMILARITY).
FT METAL 48 48 COPPER (BY SIMILARITY).
FT METAL 63 63 COPPER AND ZINC (BY SIMILARITY).
FT METAL 71 71 ZINC (BY SIMILARITY).
FT METAL 80 80 ZINC (BY SIMILARITY).
FT METAL 83 83 ZINC (BY SIMILARITY).
FT METAL 120 120 COPPER (BY SIMILARITY).
FT DISULFID 57 146 BY SIMILARITY.
SQ SEQUENCE 153 AA; 15780 MW; 1EF987CE7A0F79A0 CRC64;

Query Match 82.0%; Score 673; DB 1; Length 153;
Best Local Similarity 83.0%; Pred. No. 1.7e-55;
Matches 127; Conservative 9; Mismatches 17; Indels 0; Gaps 0;

QY 2 ATKAVCVLKGDPVQGIINFQKESNGPVKVMGSIKGLTEGLHGFHVEFGDNTAGCTSA 61
Db 1 AMKAVCVLKGDPVQGIHFEKASGEVPSQITGLTEGEHGFHVGQDNTQGCCTA 60

QY 62 GFHFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSTEDSVLSGDHCHIIIGRTLWVH 121
Db 61 GFHFNPLSRKHGPKDEERHVGDLGNVAAGKGVANVSTEDSVLSGSHSIIGRTMVVH 120

QY 122 EKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
Db 121 EKQDDLKGGNEESTKTGNAGSRLACGVIGIAQ 153

RESULT 5
SODC_HORSE
ID SODC_HORSE STANDARD; PRT; 153 AA.
AC P00443;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Superoxide dismutase [Cu-Zn] (EC 1.15.1.1).
GN SOD1.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
RN NCBI_TaxID=9796;
RP SEQUENCE FROM N.A.
RX MEDLINE=97080551; PubMed=8921896;
RA de la Rue-Domenech R., Wiedmann M., Mohammed H.O., Cummings J.F.,
RA Divers T.J., Batt C.A.;
RT "Equine motor neuron disease is not linked to Cu/Zn superoxide
RT dismutase mutations: sequence analysis of the equine Cu/Zn superoxide
RT dismutase cDNA.";
RL Gene 178:83-88(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=Testis;
RC MEDLINE=99261591; PubMed=1031206;
RA Ishida N., Katayama Y., Sato F., Hasegawa T., Mukoyama H.;
RT "The cDNA sequences of equine antioxidant enzyme genes Cu/Zn-SOD and
RT Mn-SOD, and these expressions in equine tissues.";
RL J. Vet. Med. Sci. 61:291-294(1999).
RN [3]
RP SEQUENCE.
RX MEDLINE=82052979; PubMed=7298616;
RA Lerch K., Ammer D.;
RT "Amino acid sequence of copper-zinc superoxide dismutase from horse
RT liver.";
```

```

RL J. Biol. Chem. 256:11545-11551(1981).
CC -!- FUNCTION: Destroys radicals which are normally produced within the
CC cells and which are toxic to biological systems.
CC -!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) -> O(2) + H(2)O(2).
CC -!- COFACTOR: Copper and zinc.
CC -!- SUBUNIT: HOMODIMER.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: BELONGS TO THE CU-ZN SUPEROXIDE DISMUTASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U38956; AAC48682.1; -.
DR EMBL; AB001692; BAA76921.1; -.
DR PIR; A00515; DSHOCZ.
DR HSP; P00442; ICBJ.
DR InterPro: IPR001424; SOD_CU_ZN.
DR Pfam; PF00080; sodcu; 1.
DR PRINTS; PR00068; CUZNDISMTASE.
DR ProDom; PD000469; SOD_CU_ZN; 1.
DR PROSITE; PS00087; SOD_CU_ZN_1; 1.
DR PROSITE; PS00332; SOD_CU_ZN_2; 1.
KW Oxidoreductase; Copper; Zinc; Acetylation.
FT INIT_MET 0 0
FT MOD_RES 1 1 ACETYLATION.
FT METAL 46 46 COPPER (BY SIMILARITY).
FT METAL 48 48 COPPER (BY SIMILARITY).
FT METAL 63 63 COPPER AND ZINC (BY SIMILARITY).
FT METAL 71 71 ZINC (BY SIMILARITY).
FT METAL 80 80 ZINC (BY SIMILARITY).
FT METAL 83 83 ZINC (BY SIMILARITY).
FT METAL 120 120 COPPER (BY SIMILARITY).
FT DISULFID 57 146 BY SIMILARITY.
SQ SEQUENCE 153 AA; 15940 MW; 202A190D0EFA0A43 CRC64;

Query Match 80.6%; Score 662; DB 1; Length 153;
Best Local Similarity 80.9%; Pred. No. 1.8e-54;
Matches 123; Conservative 14; Mismatches 15; Indels 0; Gaps 0;

QY 2 ATKAVCVLKGDPVQGIINFQKESNGPVKVMGSIKGLTEGLHGFHVEFGDNTAGCTSA 61
Db 1 ALKAVCVLKGDPVHGFHFEQQEQGGPVVLKGFIEGLTKGDHGFHVEFGDNTQGCCTA 60

QY 62 GFHFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSTEDSVLSGDHCHIIIGRTLWVH 121
Db 61 GAHFNPLSRKHGPKDEERHVGDLGNVTADENKADVDKSDSVLSGKHSIIGRTMVVH 120

QY 122 EKADDLGKGGNEESTKTGNAGSRLACGVIGIA 153
Db 121 EKQDDLKGGNEESTKTGNAGSRLACGVIGIA 152

RESULT 6
SODC_RABBIT
ID SODC_RABBIT STANDARD; PRT; 152 AA.
AC P09212;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Superoxide dismutase [Cu-Zn] (EC 1.15.1.1).
GN SOD1.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=New Zealand white; TISSUE=Lung;
```

```

RA Jackson R.M., Ho Y.;
RL Submitted (APR-1993) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE.
RX MEDLINE=89105317; PubMed=3214553;
RA Reinecke K., Wolf B., Michelson A.M., Puget K., Steffens G.J.,
RA Flohe L.;
RL "The amino-acid sequence of rabbit Cu-Zn superoxide dismutase.";
RL Biol. Chem. Hoppe-Seyler 369:715-725(1988).
CC -!- FUNCTION: Destroys radicals which are normally produced within the
CC cells and which are toxic to biological systems.
CC -!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -!- COFACTOR: Copper and zinc.
CC -!- SUBUNIT: HOMODIMER.
CC -!- SIMILARITY: BELONGS TO THE CU-ZN SUPEROXIDE DISMUTASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z22644; CAAB0357.1; -
DR PIR; S01134; S01134.
DR PIR; S33162; S33162.
DR HSP; P00441; 4SOD.
DR InterPro; IPR001424; SOD_CU_ZN.
DR Pfam; PF00080; sodcu; 1.
DR PRINTS; PR00068; CUZNDISMUTASE.
DR PRODOM; PD000469; SOD_CU_ZN; 1.
DR PROSITE; PS00087; SOD_CU_ZN_1; FALSE_NEG.
DR PROSITE; PS00332; SOD_CU_ZN_2; 1.
KW Oxidoreductase; Copper; Zinc; Acetylation.
FT INIT_MET 0 0
FT MOD_RES 1 1 ACETYLATION.
FT METAL 45 45 COPPER (BY SIMILARITY).
FT METAL 47 47 COPPER (BY SIMILARITY).
FT METAL 62 62 COPPER AND ZINC (BY SIMILARITY).
FT METAL 70 70 ZINC (BY SIMILARITY).
FT METAL 79 79 ZINC (BY SIMILARITY).
FT METAL 82 82 ZINC (BY SIMILARITY).
FT METAL 119 119 COPPER (BY SIMILARITY).
FT METAL 145 145 BY SIMILARITY.
FT DISULFID 56 16 A -> G (IN REF. 2).
FT CONFLICT 16 16 E -> G (IN REF. 2).
FT CONFLICT 43 43 R -> T (IN REF. 2).
FT CONFLICT 53 53 S -> A (IN REF. 2).
FT CONFLICT 151 151
SQ SEQUENCE 152 AA; 15687 MW; 501415FFCB95571D CRC64;

Query Match 79.7%; Score 654.5; DB 1; Length 152;
Best Local Similarity 82.9%; Pred. No. 8.7e-54;
Matches 126; Conservative 8; Mismatches 17; Indels 1; Gaps 1;

QY 2 ATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVEFGDNTAGCTSA 61
DB 1 ATKAVCVLKGDPVQGIINFEQK-ETIHEQK-GTGPVVVKGRITGLTEGLHGFHVEFGDNRQCTSA 59
QY 62 GPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSDVSIUSDGHCIIIGRTLTVVH 121
DB 60 GPHFNPLSRKHGGPKDEERHVGDLGNVTAGSNADVADLVIEDSVISLSDGMSVIGRTLTVVH 119
QY 122 EKADDLKGGNEESTKTGNAGSLRACGVIGIA 153
DB 120 EKEDDLKGGNDEESTKTGNAGSLRACGVIGIS 151

RESULT 7
SODC_CEREL STANDARD; PRT; 151 AA.
ID SODC_CEREL
AC O46412; O46413;

```

```

DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Superoxide dismutase [Cu-Zn] (EC 1.15.1.1).
GN SOD1.
OS Cervus elaphus (Red deer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervoidae;
OC Cervidae; Cervinae; Cervus.
OX NCBI_TaxID=9860;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=North American wapiti, and European red deer;
RX MEDLINE=98005640; PubMed=9345726;
RA He K., Wilton S.D., Tate M.L., Murphy M.P.;
RT "Characterization of the erythrocyte superoxide dismutase allozymes
RT in the deer Cervus elaphus.";
RL Anim. Genet. 28:299-301(1997).
CC -!- FUNCTION: Destroys radicals which are normally produced within the
CC cells and which are toxic to biological systems.
CC -!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -!- COFACTOR: Copper and zinc.
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: BELONGS TO THE CU-ZN SUPEROXIDE DISMUTASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U93268; AAB88115.1; -
DR EMBL; U93269; AAB88116.1; -
DR HSP; P00442; ICBJ.
DR InterPro; IPR001424; SOD_CU_ZN.
DR Pfam; PF00080; sodcu; 1.
DR PRINTS; PR00068; CUZNDISMUTASE.
DR PRODOM; PD000469; SOD_CU_ZN; 1.
DR PROSITE; PS00087; SOD_CU_ZN_1; 1.
DR PROSITE; PS00332; SOD_CU_ZN_2; FALSE_NEG.
KW Oxidoreductase; Copper; Zinc; Acetylation; Polymorphism.
FT INIT_MET 0 0
FT MOD_RES 1 1 ACETYLATION (BY SIMILARITY).
FT METAL 44 44 COPPER (BY SIMILARITY).
FT METAL 46 46 COPPER (BY SIMILARITY).
FT METAL 61 61 COPPER (BY SIMILARITY).
FT METAL 69 69 COPPER AND ZINC (BY SIMILARITY).
FT METAL 78 78 ZINC (BY SIMILARITY).
FT METAL 81 81 ZINC (BY SIMILARITY).
FT METAL 118 118 COPPER (BY SIMILARITY).
FT DISULFID 55 144 BY SIMILARITY.
FT VARIANT 8 8 L -> M (IN EUROPEAN RED DEER).
FT VARIANT 25 25 H -> N (IN EUROPEAN RED DEER).
SQ SEQUENCE 151 AA; 15697 MW; E718D474D3A22B57 CRC64;

Query Match 79.3%; Score 651; DB 1; Length 151;
Best Local Similarity 82.4%; Pred. No. 1.8e-53;
Matches 126; Conservative 8; Mismatches 17; Indels 2; Gaps 1;

QY 2 ATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVEFGDNTAGCTSA 61
DB 1 ATKAVCVLKGDPVQGIINFEAK--GHTVWVTGTSITGLTEGDHGFHVEFGDNTAGCTSA 58
QY 62 GPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSDVSIUSDGHCIIIGRTLTVVH 121
DB 59 GPHFNPLSRKHGGPKDEERHVGDLGNVTADKNGVAKVDIVDSLSLSGSEHSIIIGRTVMVH 118
QY 122 EKADDLKGGNEESTKTGNAGSLRACGVIGIA 154
DB 119 EKPDLLGRGCGNEESTKTGNARNRLACGVIGIA 151

```

```

RESULT 8
SODC_PIG
ID SODC_PIG STANDARD; PRT; 152 AA.
AC P04178;
DT 20-MAR-1987 (Rel. 04, Created)
DT 20-MAR-1987 (Rel. 04, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Superoxide dismutase [Cu-Zn] (EC 1.15.1.1).
GN SOD1.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE.
RX MEDLINE=85231234; PubMed=3891411;
RA Schinina M.E., Barra D., Simmaco M., Bossa F., Rotilio G.;
RT "Primary structure of porcine Cu,Zn superoxide dismutase.";
RL FEBS Lett. 186:267-270(1985).
CC -1- FUNCTION: Destroys radicals which are normally produced within the
CC cells and which are toxic to biological systems.
CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -1- COFACTOR: Copper and zinc.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE CU-ZN SUPEROXIDE DISMUTASE FAMILY.
DR PIR; A00514; DSPGZ.
DR HSP; P00442; ICBU.
DR InterPro: IPR001424; SOD_CU_ZN.
DR Pfam: PF00080; sodcu; 1.
DR PRINTS; P000068; CUZNDISMTASE.
DR ProDom: PD000469; SOD_CU_ZN; 1.
DR PROSITE; PS00087; SOD_CU_ZN_1; 1.
DR PROSITE; PS00332; SOD_CU_ZN_2; 1.
KW Oxidoreductase; Copper; Zinc; Acetylation.
FT MOD_RES 1 1 ACETYLATION.
FT METAL 45 45 COPPER (BY SIMILARITY).
FT METAL 47 47 COPPER (BY SIMILARITY).
FT METAL 62 62 COPPER AND ZINC (BY SIMILARITY).
FT METAL 70 70 ZINC (BY SIMILARITY).
FT METAL 79 79 ZINC (BY SIMILARITY).
FT METAL 82 82 ZINC (BY SIMILARITY).
FT METAL 119 119 COPPER (BY SIMILARITY).
FT DISULFID 56 145
FT ACT_SITE 142 142 ESSENTIAL ARGININE.
FT SEQUENCE 152 AA; 15761 MW; E60BF8C3446DEACC CRC64;

Query Match 79.2%; Score 650.5; DB 1; Length 152;
Best Local Similarity 82.4%; Pred. No. 2e-53;
Matches 126; Conservative 7; Mismatches 19; Indels 1; Gaps 1;

QY 2 ATKAVCVLKGDPVQGIINFEQKESNPVKVNGSIKGLTEGLHGFVHFEGDNTAGCTSA 61
Db 1 ATKAVCVLKGDPVQGIITVFELK-GEKTVLVGTGKLAEGDGHGFVHFQFGDNTQGCISA 59
QY 62 GPHFNPFSRKHGKGPDEERHVGDLGNVADKDGADVSTEDSVLSGDHCIIIGRTLVVH 121
Db 60 GPHFNPFSRKHGKGPDEERHVGDLGNVADKDGADVSTEDSVLSGDHCIIIGRTLVVH 119
QY 122 EKADLLGKGNNEESTKTNAGSRLACGVIGIAQ 154
Db 120 EKPPDLGRGNEESTKTNAGSRLACGVIGITQ 152

RESULT 9
SODC_BOVIN
ID SODC_BOVIN STANDARD; PRT; 151 AA.
AC P00442;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)

```

```

DE Superoxide dismutase [Cu-Zn] (EC 1.15.1.1).
GN SOD1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91088332; PubMed=2263495;
RA Gibbs L.S., Shaffer J.B.;
RT "Nucleotide sequence of bovine copper/zinc superoxide dismutase cDNA
RL generated by the polymerase chain reaction.";
RL Nucleic Acids Res. 18:7171-7171(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=75060425; PubMed=4279916;
RA Steinman H.M., Naik V.R., Abernethy J.L., Hill R.L.;
RT "Bovine erythrocyte superoxide dismutase. Complete amino acid
RL sequence.";
RL J. Biol. Chem. 249:7326-7338(1974).
RN [4]
RP DISULFIDE BOND.
RX MEDLINE=75060426; PubMed=4436313;
RA Abernethy J.L., Steinman H.M., Hill R.L.;
RT "Bovine erythrocyte superoxide dismutase. Subunit structure and
RL sequence location of the intrasubunit disulfide bond.";
RL J. Biol. Chem. 249:7339-7347(1974).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX MEDLINE=83085795; PubMed=7175933;
RA Tainer J.A., Getzoff E.D., Beem K.M., Richardson J.S.,
RA Richardson D.C.;
RT "Determination and analysis of the 2 A-structure of copper, zinc
RL superoxide dismutase.";
RL J. Mol. Biol. 160:181-217(1982).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).
RX MEDLINE=75158289; PubMed=1055410;
RA Richardson J.S., Thomas K.A., Rubin B.H., Richardson D.C.;
RT "Crystal structure of bovine Cu,Zn superoxide dismutase at 3-A
RL resolution: chain tracing and metal ligands.";
RL Proc. Natl. Acad. Sci. U.S.A. 72:1349-1353(1975).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX MEDLINE=92318263; PubMed=1619651;
RA Djinovic K., Coda A., Antolini L., Pelosi G., Desideri A., Falconi M.,
RA Rotilio G., Bolognesi M.;
RT "Crystal structure solution and refinement of the semisynthetic
RL cobalt-substituted bovine erythrocyte superoxide dismutase at 2.0-A
RN resolution.";
RN J. Mol. Biol. 226:227-238(1992).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX MEDLINE=95371126; PubMed=7643403;
RA Rypniewski W.R., Mangani S., Bruni B., Orioli P.L., Casati M.,
RA Wilson K.S.;
RT "Crystal structure of reduced bovine erythrocyte superoxide dismutase
RL at 1.9-A resolution.";
RL J. Mol. Biol. 251:282-296(1995).
RN [9]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RX MEDLINE=99194860; PubMed=10092461;
RA Hough M.A., Hasnain S.S.;
RT "Crystallographic structures of bovine copper-zinc superoxide
RL dismutase reveal asymmetry in two subunits: functionally important
RT three and five coordinate copper sites captured in the same crystal.";
RL J. Mol. Biol. 287:579-592(1999).

```


SQ SEQUENCE 151 AA; 15563 MW; EDF854EB2B17DD98 CRC64;
 Query Match 77.3%; Score 635; DB 1; Length 151;
 Best Local Similarity 80.9%; Pred. No. 5.5e-52;
 Matches 123; Conservative 8; Mismatches 19; Indels 2; Gaps 1;
 QY 2 ATKAVCLKGDGPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVEFGDNTAGCTSA 61
 DB 1 ATKAVCLKGDGPVQGTIRFEAK--GDKVVVTGTSITGLTEGDHGFHVFQDNTGCTSA 58
 QY 62 GFHFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSTEDSVISLGDHCHIIIGRTLVLVH 121
 DB 59 GFHFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSTEDSVISLGDHCHIIIGRTLVLVH 118
 QY 122 EKADDLGKGNNESTKTGNAGSLACGVIGIA 153
 DB 119 EKPDLLGRGNEESTKTGNAGSLACGVIGIA 150
 RESULT 11
 SODC_CHICK
 ID SODC_CHICK STANDARD; PRT; 153 AA.
 AC P80566; Q92059;
 DT 01-REB-1996 (Rel. 33, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Superoxide dismutase [Cu-Zn] (EC 1.15.1.1).
 GN SOD1.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Spinal cord;
 RX MEDLINE=97142141; PubMed=8988375;
 RA Stanton J.L., Wilton S.D., Laing N.G.;
 RT "Characterisation of the chicken Cu,Zn superoxide dismutase gene.";
 RL DNA Seq. 6:357-360(1996).
 RN [2]
 RP SEQUENCE.
 RC TISSUE=Erythrocyte;
 RX MEDLINE=96215439; PubMed=8647082;
 RA Schinina M.E., Carlini P., Politicelli F., Zappacosta F., Bossa F.,
 RA Calabrese L.;
 RT "Amino acid sequence of chicken Cu, Zn-containing superoxide
 RT dismutase and identification of glutathionyl adducts at exposed
 RT cysteine residues.";
 RL Eur. J. Biochem. 237:433-439(1996).
 CC -|- FUNCTION: Destroys radicals which are normally produced within the
 CC cells and which are toxic to biological systems.
 CC -|- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
 CC -|- COFACTOR: Copper and zinc.
 CC -|- SUBUNIT: HOMODIMER.
 CC -|- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -|- MASS SPECTROMETRY: MW=15600; MW_ERR=2; METHOD=Electrospray.
 CC -|- SIMILARITY: BELONGS TO THE CU-ZN SUPEROXIDE DISMUTASE FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/annouce/>
 CC or send an email to license@sib-sib.ch).

 CC EMBL: U28407; AAB88059.1; .
 CC HSP: P00441; 1SPD.
 CC InterPro: IPR001424; SOD_CU_ZN.
 CC Pfam: PF00080; sodcu; 1.
 CC PRINTS: PR00068; CUZNDISMUTASE.
 CC ProDom: PD000469; SOD_CU_ZN; 1.

DR PROSITE; PS00087; SOD_CU_ZN_1; 1.
 DR PROSITE; PS00332; SOD_CU_ZN_2; 1.
 KW Oxidoreductase; Copper; Zinc; Acetylation.
 FT INIT_MET 0 0
 FT MOD_RES 1 1 ACETYLATION.
 FT METAL 46 46 COPPER (BY SIMILARITY).
 FT METAL 48 48 COPPER (BY SIMILARITY).
 FT METAL 63 63 COPPER AND ZINC (BY SIMILARITY).
 FT METAL 71 71 ZINC (BY SIMILARITY).
 FT METAL 80 80 ZINC (BY SIMILARITY).
 FT METAL 83 83 ZINC (BY SIMILARITY).
 FT METAL 119 119 COPPER (BY SIMILARITY).
 FT METAL 153 153 BY SIMILARITY.
 FT DISULFID 57 145 WITH GLUTATHIONE.
 FT CONFLICT 13 13 A -> G (IN REF. 2).
 SQ SEQUENCE 153 AA; 15572 MW; 594F5242EFAFB7 CRC64;
 Query Match 72.2%; Score 593; DB 1; Length 153;
 Best Local Similarity 73.5%; Pred. No. 4.3e-48;
 Matches 111; Conservative 19; Mismatches 19; Indels 2; Gaps 2;
 QY 4 KAVCVLKGDGPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVEFGDNTAGCTSA 63
 DB 4 KAVCVLMKGDGPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVEFGDNTAGCTSA 62
 QY 64 HFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSTEDSVISLGDHCHIIIGRTLVLVH 123
 DB 63 HFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSTEDSVISLGDHCHIIIGRTLVLVH 121
 QY 124 ADDLKGKGNNESTKTGNAGSLACGVIGIA 154
 DB 122 SDDLGRGSDNESKLTGNAGSLACGVIGIA 152
 RESULT 12
 SODC_CARCR
 ID SODC_CARCR STANDARD; PRT; 166 AA.
 AC P80174;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Superoxide dismutase [Cu-Zn] (EC 1.15.1.1).
 OS Caretta caretta (Loggerhead).
 OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
 OC Testudines; Cryptodira; Chelonioidea; Chelonidae; Caretta.
 OX NCBI_TaxID=8467;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Liver;
 RX MEDLINE=93170321; PubMed=8436140;
 RA Schinina M.E., Bossa F., Lania A., Capo C.R., Carlini P.,
 RA Calabrese L.;
 RT "The primary structure of turtle Cu,Zn superoxide dismutase.
 RT Structural and functional irrelevance of an insert conferring
 RT proteolytic susceptibility.";
 RL Eur. J. Biochem. 211:843-849(1993).
 CC -|- FUNCTION: Destroys radicals which are normally produced within the
 CC cells and which are toxic to biological systems.
 CC -|- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
 CC -|- COFACTOR: Copper and zinc.
 CC -|- SUBUNIT: HOMODIMER.
 CC -|- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -|- SIMILARITY: BELONGS TO THE CU-ZN SUPEROXIDE DISMUTASE FAMILY.
 DR PIR: S29782; S29782.
 DR HSP: P00441; 4SOD.
 DR InterPro: IPR001424; SOD_CU_ZN.
 DR Pfam: PF00080; sodcu; 1.
 DR PRINTS: PR00068; CUZNDISMUTASE.
 DR ProDom: PD000469; SOD_CU_ZN; 1.
 DR PROSITE; PS00087; SOD_CU_ZN_1; 1.
 DR PROSITE; PS00332; SOD_CU_ZN_2; 1.
 KW Oxidoreductase; Copper; Zinc; Repeat.
 FT MOD_RES 1 1 BLOCKED.

```

FT REPEAT 14 17
FT REPEAT 18 21
FT REPEAT 22 25
FT REPEAT 26 29
FT METAL 58 58
FT METAL 60 60
FT METAL 75 75
FT METAL 83 83
FT METAL 92 92
FT METAL 95 95
FT METAL 132 132
FT METAL 158 158
FT DISULFID 69 158
SQ SEQUENCE 166 AA; 17146 MW; 84306991A0299870 CRC64;

Query Match 71.8%; Score 589.5; DB 1; Length 166;
Best Local Similarity 69.9%; Pred. No. 9.9e-48;
Matches 114; Conservative 16; Mismatches 20; Indels 13; Gaps 2;

QY 4 KAVCVLKGDPVKEPVKGPVKPVKGIYFEQ-QCGNPGVTLSGSIITGLTEGKHGFHVHEF 51
DQ 4 KAVCVLKGDPVKEPVKGPVKPVKGIYFEQ-QCGNPGVTLSGSIITGLTEGKHGFHVHEF 62

QY 52 GONTAGCTAGAHFNPPLSKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGH 111
DQ 63 GONTAGCTAGAHFNPPLSKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGH 122

QY 112 CIIGRTLVVHERKADDLGKGGNEESTKTGNAGSLACGVIGIAQ 154
DQ 123 SIIGRTLVVHERKADDLGKGGNEESTKTGNAGSLACGVIGIAQ 165

RESULT 13
SODC_BRARE
ID SODC_BRARE STANDARD; PRT; 154 AA.
AC 073872:
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Superoxide dismutase [Cu-Zn] (EC 1.15.1.1).
GN SOD1 OR CUZD.
OS Brachydanio rerio (zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Ken C.F., Lin C.T.;
RT "Cloning and characterization of a cDNA for Cu/Zn-superoxide dismutase
from zebrafish."
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Destroys radicals which are normally produced within the
cells and which are toxic to biological systems.
CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -1- COFACTOR: Copper and zinc.
CC -1- SUBUNIT: HOMODIMER (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE CU-ZN SUPEROXIDE DISMUTASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DQ EMBL; Y12236; CAA72925.1; -.
DQ ZFIN; ZDB-GENE-990415-258; cuzn.
DQ InterPro; IPR001424; SOD_CU_ZN.
DQ Pfam; PF00080; sodcu; 1.
DQ PRINTS; PR00068; CUZNDISMUTASE.

```

```

DR PRODOM; PD000469; SOD_CU_ZN; 1.
DR PROSITE; PS00087; SOD_CU_ZN_1; 1.
DR PROSITE; PS00332; SOD_CU_ZN_2; 1.
KW Oxidoreductase; Copper; Zinc.
FT METAL 47 47 COPPER (BY SIMILARITY).
FT METAL 49 49 COPPER (BY SIMILARITY).
FT METAL 64 64 COPPER AND ZINC (BY SIMILARITY).
FT METAL 72 72 ZINC (BY SIMILARITY).
FT METAL 81 81 ZINC (BY SIMILARITY).
FT METAL 84 84 ZINC (BY SIMILARITY).
FT METAL 121 121 COPPER (BY SIMILARITY).
FT METAL 147 147 BY SIMILARITY.
SQ SEQUENCE 154 AA; 15953 MW; 62F44CAL46C5057C CRC64;

Query Match 70.5%; Score 579; DB 1; Length 154;
Best Local Similarity 70.1%; Pred. No. 8.6e-47;
Matches 108; Conservative 10; Mismatches 36; Indels 0; Gaps 0;

QY 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVHEFGDNTAGCTS 60
DQ 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVHEFGDNTAGCTS 60

QY 61 AGPHFNPLSKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIGRTLVV 120
DQ 61 AGPHFNPLSKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIGRTLVV 120

QY 121 HEKADDLGKGGNEESTKTGNAGSLACGVIGIAQ 154
DQ 121 HEKADDLGKGGNEESTKTGNAGSLACGVIGIAQ 154

RESULT 14
SODC_PRIGL
ID SODC_PRIGL STANDARD; PRT; 152 AA.
AC P11418:
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Superoxide dismutase [Cu-Zn] (EC 1.15.1.1).
OS Prionace glauca (Blue shark).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphi; Galeoidei; Carcharhiniformes;
OC Carcharhinidae; Prionace.
OX NCBI_TaxID=7815;
RN [1]
RP SEQUENCE.
RA MEDLINE-89290032; PubMed-2500367;
RA Calabrese L., Politicelli F., O'Neill P., Galtieri A., Barra D.,
RA Schinina M.E., Bossa F.;
RT "Substitution of arginine for lysine 134 alters electrostatic
parameters of the active site in shark Cu,Zn superoxide dismutase."
RL FEBS Lett. 250:49-52(1989).
CC -1- FUNCTION: Destroys radicals which are normally produced within the
cells and which are toxic to biological systems.
CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -1- COFACTOR: Copper and zinc.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE CU-ZN SUPEROXIDE DISMUTASE FAMILY.
DQ PIR; S04623; S04623.
DQ PIR; S45643; S45643.
DQ HSSP; P00442; ICBJ.
DQ InterPro; IPR001424; SOD_CU_ZN.
DQ Pfam; PF00080; sodcu; 1.
DQ PRINTS; PR00068; CUZNDISMUTASE.
DQ PRODOM; PD000469; SOD_CU_ZN; 1.
DQ PROSITE; PS00087; SOD_CU_ZN_1; 1.
DQ PROSITE; PS00332; SOD_CU_ZN_2; 1.
KW Oxidoreductase; Copper; Zinc.
FT METAL 44 44 COPPER.
FT METAL 46 46 COPPER.
FT METAL 61 61 COPPER AND ZINC.
FT METAL 69 69 ZINC.

```

```
FT METAL 78 78 ZINC.
FT METAL 81 81 ZINC.
FT METAL 118 118 COPPER.
FT DISULFID 55 144 BY SIMILARITY.
SQ SEQUENCE 152 AA; 15840 MW; 66176424A4F23C5AE CRC64;

Query Match 68.0%; Score 558.5; DB 1; Length 152;
Best Local Similarity 70.9%; Pred. No. 6.7e-45;
Matches 107; Conservative 15; Mismatches 28; Indels 1; Gaps 1;

QY 4 KAVCVLKGDPVQGIINFEQKESGPKVWGSIKGLTEGLHGFHVEFGDNTAGCTSAGP 63
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2 KAVCVLKGTEGTVLPQ-AADPVTGLKSGITGLTKGKGFHVAFGDNTAGCTSAGP 60
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 64 HFNPLSRKHGPKDERHVGDLGNVTADKDGADVSDISGSDHCIIIGRTLVVHEK 123
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 HYNPFKNGHGGPDEERHVGDLGNVTADKDGADVSDISGSDHCIIIGRTLVVHEK 120
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 124 ADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
Db 121 EDDLKGGDEESLRYGNAGSRLACGVIGIAK 151

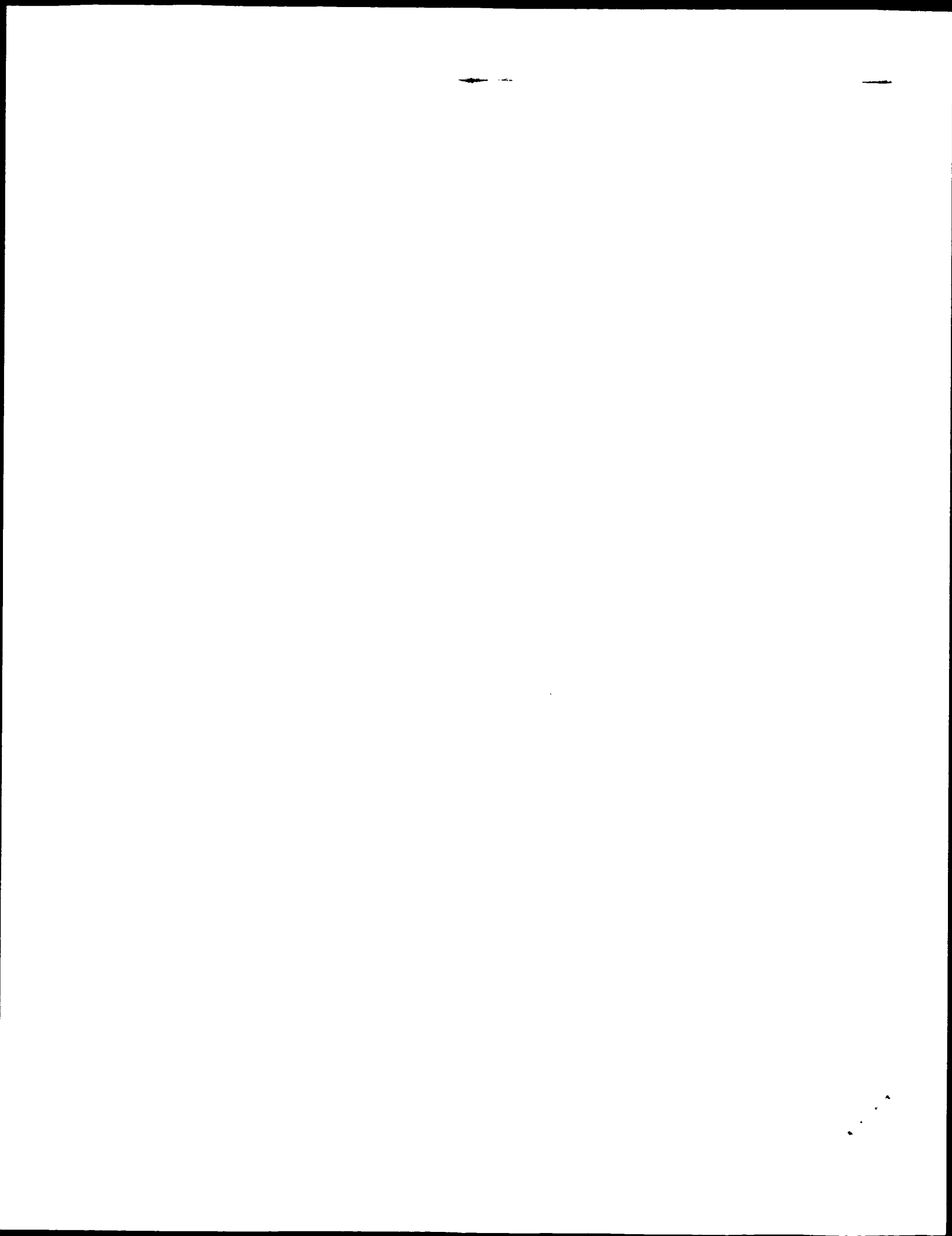
RESULT 15
SODD.XENLA
ID SODD.XENLA STANDARD; PRT; 150 AA.
AC P15107;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Superoxide dismutase [Cu-Zn] 2 (EC 1.15.1.1) (xSODB).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90221905; PubMed=2326205;
RA Carri M.T., Battistoni A., Mariottini P., Rotilio G.;
RT "Xenopus laevis Cu,Zn superoxide dismutase B cDNA sequence.";
RL Nucleic Acids Res. 18:1641-1641(1990).
RN [2]
RP SEQUENCE.
RX MEDLINE=89321563; PubMed=2751312;
RA Schinina M.E., Barra D., Bossa F., Calabrese L., Montesano L.,
RA Carri M.T., Mariottini P., Amaldi F., Rotilio G.;
RT "Primary structure from amino acid and cDNA sequences of two Cu,Zn
RT superoxide dismutase variants from Xenopus laevis.";
RL Arch. Biochem. Biophys. 272:507-515(1989).
RN [3]
RP 3D-STRUCTURE MODELING.
RX MEDLINE=91376038; PubMed=1896428;
RA Falconi M., Rotilio G., Desideri A.;
RT "Modelling the three-dimensional structure and electrostatic
RT potential field of the two Cu,Zn superoxide dismutase variants from
RT Xenopus laevis.";
RL Proteins 10:149-155(1991).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (1.49 ANGSTROMS).
RX Djinovic Carugo K., Battistoni A., Carri M.T., Polticelli F.,
RA Desideri A., Rotilio G., Coda A., Wilson K.S., Bolognesi M.;
RT "Three-dimensional structure of Xenopus laevis Cu,Zn superoxide
RT dismutase b. determined by X-ray crystallography at 1.5-A
RT resolution.";
RL Acta Crystallogr. D 52:176-188(1996).
CC CC
CC -1- FUNCTION: Destroys radicals which are normally produced within the
CC cells and which are toxic to biological systems.
CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -1- COFACTOR: Copper and zinc.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE CU-ZN SUPEROXIDE DISMUTASE FAMILY.
```

```
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X51518; CAA35890.1; -.
DR PIR: S05022; S05022.
DR PIR: S09568; S09568.
DR PDB: 1XSO; 10-JUL-95.
DR InterPro: IPR001424; SOD_CU_ZN.
DR Pfam: PF00080; sodcu; 1.
DR PRINTS: PR00068; CUZNDISMTASE.
DR PRODOM: PD000469; SOD_CU_ZN; 1.
DR PROSITE: PS00087; SOD_CU_ZN_1; 1.
DR PROSITE: PS00332; SOD_CU_ZN_2; 1.
KW Oxidoreductase; Copper; zinc; 3D-structure.
FT INIT_MET 0 0
FT METAL 44 44 COPPER (BY SIMILARITY).
FT METAL 46 46 COPPER (BY SIMILARITY).
FT METAL 61 61 COPPER AND ZINC (BY SIMILARITY).
FT METAL 69 69 ZINC (BY SIMILARITY).
FT METAL 78 78 ZINC (BY SIMILARITY).
FT METAL 81 81 ZINC (BY SIMILARITY).
FT METAL 117 117 COPPER (BY SIMILARITY).
FT DISULFID 55 143 BY SIMILARITY.
FT CONFLICT 60 60 P -> S (IN REF. 1).
SQ SEQUENCE 150 AA; 15297 MW; 0F5236CEEB236250 CRC64;

Query Match 65.9%; Score 541; DB 1; Length 150;
Best Local Similarity 69.6%; Pred. No. 2.7e-43;
Matches 103; Conservative 16; Mismatches 27; Indels 2; Gaps 2;

QY 4 KAVCVLKGDPVQGIINFEQKESGPKVWGSIKGLTEGLHGFHVEFGDNTAGCTSAGP 63
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2 KAVCVLAGSDGVKGVHFEQD-EGAVSVECKIEGLTDGLHGFHVEFGDNTAGCTSAGP 60
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 64 HFNPLSRKHGPKDERHVGDLGNVTADKDGADVSDISGSDHCIIIGRTLVVHEK 123
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 HFNPNKNHGA PGDTRHVGDLGNVTAE-GGVAQFKITDSLISLKGPNSSIIGRTAVVHEK 119
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 124 ADDLGKGGNEESTKTGNAGSRLACGVIG 151
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 120 ADDLGKGGNDESLKTGNAGSRLACGVIG 147
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Search completed: January 28, 2003, 14:15:18
Job time : 3.35564 secs
```



GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:04:57 ; Search time 7.78203 Seconds
(without alignments)
4077.505 Million cell updates/sec

Title: US-09-904-987-6

Perfect score: 821

Sequence: 1 MATKAVCVLKGDGPVQGIIN.....STKTGNAGSLRACGVIGIAQ 154

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
1	807	98.3	179	3	Q12728	Q12728 Schizosacch
2	638.5	77.8	153	6	Q8WNN6	Q8WNN6 canis famil
3	632.5	77.0	147	6	Q95ME5	Q95ME5 sus scrofa
4	563	68.6	154	13	Q8QHI0	Q8QHI0 oncorhynch
5	508.5	61.9	154	3	Q9C0N4	Q9C0N4 filobasidie
6	503.5	61.3	154	3	Q9C0T2	Q9C0T2 cryptococcu
7	500	60.9	153	5	Q9U4X3	Q9U4X3 drosophila
8	499	60.8	153	5	Q9U4X5	Q9U4X5 drosophila
9	496	60.4	153	5	Q9U4X2	Q9U4X2 drosophila
10	489.5	59.6	154	3	Q9C0S4	Q9C0S4 cryptococcu
11	489	59.6	145	5	Q96348	Q96348 drosophila
12	489	59.6	153	5	Q9U4X4	Q9U4X4 drosophila
13	479	58.3	158	10	Q9AR76	Q9AR76 populus tre
14	478.5	58.3	158	3	Q96WH8	Q96WH8 debaryomyce
15	477	58.1	145	5	O16861	O16861 chymomyza p
16	477	58.1	145	5	O96353	O96353 drosophila

17	476	58.0	145	5	O96349	O96349 drosophila
18	475	57.9	145	5	O16863	O16863 zaprionus t
19	473.5	57.7	183	3	O9C402	O9C402 cryptococcu
20	473.5	57.7	215	10	Q944B7	Q944B7 pinus pinas
21	472	57.5	158	10	Q9AR77	Q9AR77 populus tre
22	471.5	57.4	151	5	O77243	O77243 dictyosteli
23	471	57.4	145	5	O27630	O27630 drosophila
24	471	57.4	145	5	O8WRP5	O8WRP5 taenia soli
25	470	57.2	152	5	O96352	O96352 drosophila
26	468.5	57.1	210	8	O9BA07	O9BA07 populus tre
27	468	57.0	145	5	O96351	O96351 drosophila
28	467.5	56.9	152	10	Q9AR78	Q9AR78 populus tre
29	465.5	56.7	201	10	O24400	O24400 triticum ae
30	465	56.6	146	5	Q27770	Q27770 drosophila
31	464	56.5	145	5	O96356	O96356 drosophila
32	462.5	56.3	201	8	O96123	O96123 triticum ae
33	461	56.2	145	5	O16862	O16862 drosophila
34	461	56.2	146	5	O9XY94	O9XY94 fasciola he
35	460.5	56.1	201	5	O18483	O18483 onchocerca
36	460.5	56.1	201	10	Q93Y61	Q93Y61 dichanthelli
37	460	56.0	145	5	Q27569	Q27569 drosophila
38	460	56.0	158	5	O9Y0A5	O9Y0A5 acanthochei
39	459.5	56.0	156	10	Q9AR80	Q9AR80 pinus sylve
40	458.5	55.8	216	10	Q9SUJ7	Q9SUJ7 arabidopsis
41	458	55.8	153	3	O8X1S8	O8X1S8 aspergillus
42	454.5	55.4	152	10	Q8S3V1	Q8S3V1 sandersonia
43	453	55.2	161	10	O48654	O48654 marchantia
44	452.5	55.1	152	10	Q9ZS54	Q9ZS54 pinus sylve
45	452.5	55.1	152	10	P93801	P93801 zea mays (m

ALIGNMENTS

RESULT 1

Q12728 PRELIMINARY; PRT; 179 AA.
AC Q12728
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Yeast HB-SOD (Fragment).
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91358420; PubMed=1885572;
RA Inoue M., Watanabe N., Matsuno K., Sasaki J., Tanaka Y.,
RA Hatanaka-Sawai H., Amachi T.;
RT "Expression of a hybrid Cu/Zn-type superoxide dismutase which has high
RT affinity for heparin-like proteoglycans on vascular endothelial
RT cells.";
RL J. Biol. Chem. 266:16409-16414(1991).
DR EMBL; D90358; BAA14373.1; -;
DR HSP; P00441; ISP.D.
DR InterPro; IPR001424; SOD_CU_2N.
DR Pfam; PF00080; sodcu; 1.
DR PRINTS; PR00068; CUZNDISMTASE.
DR ProDom; PD000469; SOD_CU_2N; 1.
DR PROSITE; PS00087; SOD_CU_2N_1; 1.
DR PROSITE; PS00332; SOD_CU_2N_2; 1.
FT NON_TER 1
FT NON_TER 179
SQ SEQUENCE 179 AA; 18995 MW; 42D662A6785F6135 CRC64;

Query Match 98.3%; Score 807; DB 3; Length 179;
Best Local Similarity 100.0%; Pred. No. 2.3e-69;
Matches 151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 ATKAVCVLKGDGPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVHFGDNTAGTCTSA 61

```

|||||
Db 1 ATRAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVHFEFGDNTAGCTSA 60
QY 62 GPHFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTVH 121
Db 61 GPHFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTVH 120
QY 122 EKADDLGKGGNEESTKTGNAGSLACGVIGI 152
Db 121 EKADDLGKGGNEESTKTGNAGSLACGVIGI 151

RESULT 2
Q8WNN6 PRELIMINARY; PRT; 153 AA.
AC Q8WNN6;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Cu/Zn superoxide dismutase.
GN SOD1.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RA Green S.L., Tolwani R.J., Varma S., Quignon P., Galibert F.,
RA Cork L.C.;
RT "The canine Cu/Zn superoxide dismutase gene (SOD1).";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF346417; AAL61608.1; -.
DR InterPro: IPR001424; SOD_CU_ZN.
DR Pfam: PF00080; soccu.1.
DR PRINTS; PR00068; CUZNDISMTASE.
DR ProDom: PD000469; SOD_CU_ZN.1.
DR PROSITE; PS00332; SOD_CU_ZN_2; UNKNOWN_1.
SQ SEQUENCE 153 AA; 15912 MW; 0D7900E59C57B6B0 CRC64;

Query Match 77.8%; Score 638.5; DB 6; Length 153;
Best Local Similarity 80.5%; Pred. No. 2.3e-53;
Matches 124; Conservative 11; Mismatches 18; Indels 1; Gaps 1;

QY 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVHFEFGDNTAGCTS 60
Db 1 MEMKAVCVLKGQGPVESTHFVQKGS-GPVVSGTITGLTEGEHGFHVHVFEDXTQGCTS 59
QY 61 AGPHFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 120
Db 60 AGPHFNPLSRKHGPKDEERHVGDLGNVTAGKGVADVSIEDSVIALSGDHSIIIGRTMVV 119
QY 121 HEKADDLGKGGNEESTKTGNAGSLACGVIGIAQ 154
Db 120 HEKADDLGKGGNEESTKTGNAGSLACGVIGIAQ 153

RESULT 3
Q95WE5 PRELIMINARY; PRT; 147 AA.
AC Q95WE5;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Superoxide dismutase 1 (Fragment).
GN SOD1.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Llu B.S., Ding S.T.;
RT "Expression of superoxide dismutase 1 in pigs.";
```

```

Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF396674; AAR84037.1; -.
DR InterPro: IPR001424; SOD_CU_ZN.
DR Pfam: PF00080; soccu.1.
DR ProDom: PD000469; SOD_CU_ZN; 1.
DR PROSITE; PS00087; SOD_CU_ZN_1; UNKNOWN_1.
DR PROSITE; PS00332; SOD_CU_ZN_2; UNKNOWN_1.
FT NON_TER 1
FT NON_TER 147
SQ SEQUENCE 147 AA; 15246 MW; 62156DD7EC5430A4 CRC64;

Query Match 77.0%; Score 632.5; DB 6; Length 147;
Best Local Similarity 82.4%; Pred. No. 8.3e-53;
Matches 122; Conservative 7; Mismatches 18; Indels 1; Gaps 1;

QY 4 KAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVHFEFGDNTAGCTSAGP 63
Db 1 KAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVHFEFGDNTAGCTSAGP 59
QY 64 HFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTVVHEK 123
Db 60 HFNPLSRKHGPKDEERHVGDLGNVTAGKGVADVSIEDSVIALSGDHSIIIGRTMVVHEK 119
QY 124 ADLGLKGGNEESTKTGNAGSLACGVIG 151
Db 120 PDLGLRGNGNEESTKTGNAGSLACGVIG 147

RESULT 4
Q8QHI0 PRELIMINARY; PRT; 154 AA.
AC Q8QHI0;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Cu/Zn superoxide dismutase.
GN SOD1.
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8022;
RN [1]
RP SEQUENCE FROM N.A.
RA Brunelli J.P., Thorgaard G.H.;
RT "Rainbow trout Cu/Zn-superoxide dismutase (SOD1) mRNA.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF469663; AAL79162.1; -.
SQ SEQUENCE 154 AA; 15755 MW; CA3C983D78067A18 CRC64;

Query Match 68.6%; Score 563; DB 13; Length 154;
Best Local Similarity 68.2%; Pred. No. 3.8e-46;
Matches 105; Conservative 17; Mismatches 32; Indels 0; Gaps 0;

QY 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVHFEFGDNTAGCTS 60
Db 1 MAMKAVCVLKGTEGTVTFEFGADGPKVLIGEISGLAFGEHGFHVHAYGDNNTGMS 60
QY 61 AGPHFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 120
Db 61 AGPHFNPHNQTHGGPTDAVRHVGDGNVTAGADNVAKINQDKMLTLGPPDSIIIGRTWVI 120
QY 121 HEKADDLGKGGNEESTKTGNAGSLACGVIGIAQ 154
Db 121 HEKADDLGKGGNEESTKTGNAGSLACGVIGIAQ 154

RESULT 5
Q9CON4 PRELIMINARY; PRT; 154 AA.
AC Q9CON4;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
```

DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Copper zinc superoxide dismutase.
GN SOD1.
OS Filobasidiella bacillisporea.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
OX NCBI_TaxID=37769;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-WM 780-C;
RA Chaturvedi S., Chaturvedi V.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-NYS 155B;
RA Chaturvedi S., Hamilton A.J., Chaturvedi V.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-NYS 155B;
RA Hamilton A.J., Chaturvedi S., Chaturvedi V.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE
CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: 2 SUPEROXIDE + 2 H(+) = O(2) + H(2)O(2).
CC -1- SIMILARITY: BELONGS TO THE CU-ZN SUPEROXIDE DISMUTASE FAMILY.
DR EMBL: AF248051; AAK31920.1; -;
DR EMBL: AF248049; AAK31918.1; -;
DR EMBL: AF248050; AAK31919.1; -;
DR HSSP: P00442; 3SOD.
DR InterPro: IPR001424; SOD_CU_ZN.
DR Pfam: PF00080; sodcu; 1.
DR PRINTS: PR00068; CUZNDISMTASE.
DR PRODOM: PD000469; SOD_CU_ZN_1.
DR PROSITE: PS00087; SOD_CU_ZN_1; 1.
DR PROSITE: PS00332; SOD_CU_ZN_2; 1.
DR Copper; Oxidoreductase; Zinc.
KW SEQUENCE 154 AA; 15637 MW; CM4C2B4BA8F43FE4 CRC64;
SQ
Query Match 61.9%; Score 508.5; DB 3; Length 154;
Best Local Similarity 64.9%; Pred. No. 6.2e-41;
Matches 98; Conservative 11; Mismatches 41; Indels 1; Gaps 1;
QY 4 KAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGL-TBGLHGFHVHFGDNTAGCTSGAG 62
DB 3 KAVAVLKGDPVGTITFTQREGAPVTGSDIKNLDAERGFHVHFGDNTAGCTSGAG 62
QY 63 PHFNPLSRKHGPKDEERHVGDLGNVTDKGVADVSDISLSGDHCHITGRTLVVH 122
DB 63 PHFNPHKNGHAPSDESRHVGDLGNVTDKGVASVNIKSLSLFGPYSIIGRTIVVHA 122
QY 123 KADDLGKGNEESTKTNAGSLACGVIGIA 153
DB 123 GTDDFGKGGNAESLKTGNAGAACGVIGIS 153
Query Match 61.9%; Score 508.5; DB 3; Length 154;
Best Local Similarity 64.9%; Pred. No. 6.2e-41;
Matches 98; Conservative 11; Mismatches 41; Indels 1; Gaps 1;
QY 4 KAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGL-TBGLHGFHVHFGDNTAGCTSGAG 62
DB 3 KAVAVLKGDPVGTITFTQREGAPVTGSDIKNLDAERGFHVHFGDNTAGCTSGAG 62
QY 63 PHFNPLSRKHGPKDEERHVGDLGNVTDKGVADVSDISLSGDHCHITGRTLVVH 122
DB 63 PHFNPHKNGHAPSDESRHVGDLGNVTDKGVASVNIKSLSLFGPYSIIGRTIVVHA 122
QY 123 KADDLGKGNEESTKTNAGSLACGVIGIA 153
DB 123 GTDDFGKGGNAESLKTGNAGAACGVIGIS 153
RESULT 6
Q9COT2 PRELIMINARY; PRT; 154 AA.
AC Q9COT2;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 20, Last annotation update)
DE Copper zinc superoxide dismutase.
GN SOD1.
OS Cryptococcus neoformans var. neoformans.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
OX NCBI_TaxID=40410;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-B-3501;
RA Chaturvedi S., Hamilton A.J., Chaturvedi V.;

Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-B-3501;
RA Chaturvedi S., Chaturvedi V.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE
CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: 2 SUPEROXIDE + 2 H(+) = O(2) + H(2)O(2).
CC -1- SIMILARITY: BELONGS TO THE CU-ZN SUPEROXIDE DISMUTASE FAMILY.
DR EMBL: AF248048; AAK31917.1; -;
DR EMBL: AF248047; AAK31916.1; -;
DR HSSP: P00441; 4SOD.
DR InterPro: IPR001424; SOD_CU_ZN.
DR Pfam: PF00080; sodcu; 1.
DR PRINTS: PR00068; CUZNDISMTASE.
DR PRODOM: PD000469; SOD_CU_ZN_1.
DR PROSITE: PS00087; SOD_CU_ZN_1; 1.
DR PROSITE: PS00332; SOD_CU_ZN_2; 1.
DR Copper; Oxidoreductase; Zinc.
KW SEQUENCE 154 AA; 15779 MW; OB0054E88FECB35 CRC64;
SQ
Query Match 61.3%; Score 503.5; DB 3; Length 154;
Best Local Similarity 66.9%; Pred. No. 1.9e-40;
Matches 101; Conservative 7; Mismatches 42; Indels 1; Gaps 1;
QY 4 KAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGL-TEGLHGFHVHFGDNTAGCTSGAG 62
DB 3 KAVAVLKGDPVGTITFTQSEGAFCVSGEIKNLDAERGFHVHFGDNTAGCTSGAG 62
QY 63 PHFNPLSRKHGPKDEERHVGDLGNVTDKGVADVSDISLSGDHCHITGRTLVVH 122
DB 63 PHFNPHKNGHGPSAAERHVGDLGNVTDKGVAVMVISLFGPHSIIGRSMVHA 122
QY 123 KADDLGKGNEESTKTNAGSLACGVIGIA 153
DB 123 GTDDLGKGGNEESLKTGNAGARLACGVIGIA 153
RESULT 7
Q9U4X3 PRELIMINARY; PRT; 153 AA.
ID Q9U4X3;
AC Q9U4X3;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE CU-ZN superoxide dismutase.
GN SOD.
OS Drosophila yakuba (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7245;
RN [1]
RP SEQUENCE FROM N.A.
RA Arxontaki K., Kastanis P., Tsakas S., Loukas M., Ellipoulos E.;
RT "Phylogenetic analysis of Drosophila melanogaster group based on Cu-Zn
RT superoxide dismutase gene sequences.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE
CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: 2 SUPEROXIDE + 2 H(+) = O(2) + H(2)O(2).
CC -1- SIMILARITY: BELONGS TO THE CU-ZN SUPEROXIDE DISMUTASE FAMILY.
DR EMBL: AF127159; AAF23598.1; -;
DR HSSP: P00441; 4SOD.
DR FlyBase: FBgn0022829; Dyak\Sod.
DR InterPro: IPR001424; SOD_CU_ZN.
DR Pfam: PF00080; sodcu; 1.
DR PRINTS: PR00068; CUZNDISMTASE.
DR PRODOM: PD000469; SOD_CU_ZN_1.
DR PROSITE: PS00087; SOD_CU_ZN_1; 1.
DR PROSITE: PS00332; SOD_CU_ZN_2; 1.
DR Copper; Oxidoreductase; Zinc.
KW

GN HIPI-SOD1.
OS Eukaryota; Tremula x Populus tremuloides.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Malpighiales; Salicaceae; Populus.
OX NCBI_TaxID=47664;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21362240; PubMed=11469593;
RA Schinkel H., Hertzberg M., Wingsle G.;
RT "A small family of novel Cu²⁺-superoxide dismutases with high
isoelectric points in hybrid aspen.";
PL Planta 213:272-279 (2001).
CC -!- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE
CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: 2 SUPEROXIDE + 2 H(+) = O(2) + H(2)O(2).
CC -!- SIMILARITY: BELONGS TO THE CU-ZN SUPEROXIDE DISMUTASE FAMILY.
DR EMBL: AJ778671; CAC33847.1; -
DR HSP; P07508; I5R0.
DR InterPro: IPR001424; SOD_CU_ZN.
DR Pfam: PF00080; sodcu; 1.
DR PRINTS: PR00068; CUZNDISMTASE.
DR ProDom: PD000469; SOD_CU_ZN; 1.
DR PROSITE: PS00087; SOD_CU_ZN_1; 1.
DR PROSITE: PS00332; SOD_CU_ZN_2; 1.
KW Copper; Oxidoreductase; Zinc.
SQ SEQUENCE 158 AA; 15976 MW; C44B29B9AFC44A70 CRG64;
Query Match 58.3%; Score 479; DB 10; Length 158;
Best Local Similarity 59.4%; Pred. No. 4.2e-38;
Matches 92; Conservative 18; Mismatches 41; Indels 4; Gaps 2;
QY 1 MAT---KAVCVLKDGDPVQGIINFEQKESNGPVKVGSIKGLTEGLGHGFHVEFGDNTAG 57
DB 1 MATGSKAVALITGDSNVGRSLHFIO-EPNGATHVTGRITGLSPGLGHGFHIALGDTTNG 59
QY 58 CTSAAPHNPLSRKHGKGPDEERHVGDLGNVTADKGVADVSDIESTSVLSGDHCHIGRT 117
DB 60 CNSTGPHNPLKDKHAPSNEHRHAGDLGNITAGSGVAEVSIKDLQIPLSGMHSILGRA 119
QY 118 LVVHEKADDLKGGNEESTKTGNAGSRLACGVIGI 152
DB 120 VVHADPDLKGGHLSKTTGNAGARVCGGIIGL 154
RESULT 14
ID Q96WH8 PRELIMINARY; PRT; 158 AA.
AC Q96WH8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Copper-zinc superoxide dismutase (EC 1.15.1.1) (Fragment).
GN SOD1.
OS Debaryomyces hansenii (Yeast) (Torulaspora hansenii).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Debaryomyces.
OX NCBI_TaxID=4959;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=CIBNOR C61;
RA Hernandez-Saavedra N.Y., Romero-Geraldo R.;
RT "Cloning and sequencing of a genomic region encoding a Cu,Zn
superoxide dismutase enzyme from several marine strains of
RT Debaryomyces (Lodder & Krieger-van Rij) genus.";
RN Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=CIBNOR C61;
RA Hernandez-Saavedra N.Y.;
RT "Universal primers to amplify the coding sequence of the enzyme
RT copper-zinc superoxide dismutase of marine yeast and yeast-like
organisms,";

RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF327448; AAK82335.1; -
DR InterPro: IPR001424; SOD_CU_ZN.
DR Pfam: PF00080; sodcu; 1.
DR ProDom: PD000469; SOD_CU_ZN; 1.
DR PROSITE: PS00087; SOD_CU_ZN_1; UNKNOWN_1.
DR PROSITE: PS00332; SOD_CU_ZN_2; UNKNOWN_1.
KW Oxidoreductase.
FT NON_TER 158
SQ SEQUENCE 158 AA; 16430 MW; 5E2122270004E367 CRG64;
Query Match 58.3%; Score 478.5; DB 3; Length 158;
Best Local Similarity 59.3%; Pred. No. 4.7e-38;
Matches 89; Conservative 19; Mismatches 41; Indels 1; Gaps 1;
QY 4 KAVCVLKDGDPVQGIINFEQKESNGPVKVGSIKGL-TEGLGHGFHVEFGDNTAGCTSG 62
DB 3 QAVAVLRGDSKVIQVGNFEQSESDPTFITWEISGNDANALRGFHVHTFGDNTNGCTSG 62
QY 63 PHFNPLSRKHGKGPDEERHVGDLGNVTADKGVADVSDIESTSVLSGDHCHIGRTLVVHE 122
DB 63 PHFNPTKEHGAPEDDNRHVGDLGNVTDTSGVAKGSKDLFVKLIGNSILGRTVVIHA 122
QY 123 KADDLKGKGGNEESTKTGNAGSRLACGVIGI 152
DB 123 GTDDLKGGNAESKKTGNAGARLACGVIGL 152
RESULT 15
ID Q16861 PRELIMINARY; PRT; 145 AA.
AC Q16861;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Superoxide dismutase (EC 1.15.1.1) (Cu-Zn) (fragment).
GN SOD.
OS Chymomyza procnevis.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Chymomyza.
OX NCBI_TaxID=59308;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94300618; PubMed=7545938;
RA Kwiatkowski J., Skarecky D., Bailey K., Ayala F.J.;
RT "Phylogeny of Drosophila and related genera inferred from the
RT nucleotide sequence of the Cu,Zn Sod gene.";
RL J. Mol. Evol. 38:443-454 (1994).
RN [2]
RP SEQUENCE FROM N.A.
RA Kwiatkowski J.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF021821; AAB80925.1; -
DR HSP; P00441; 4SOD.
DR Flybase: FBgn0021739; Cpro\Sod.
DR InterPro: IPR001424; SOD_CU_ZN.
DR Pfam: PF00080; sodcu; 1.
DR PRINTS: PR00068; CUZNDISMTASE.
DR ProDom: PD000469; SOD_CU_ZN; 1.
DR PROSITE: PS00087; SOD_CU_ZN_1; 1.
DR PROSITE: PS00332; SOD_CU_ZN_2; 1.
FT NON_TER 1
SQ SEQUENCE 145 AA; 14999 MW; BA201B66C3B77A33 CRG64;
Query Match 58.1%; Score 477; DB 5; Length 145;
Best Local Similarity 60.1%; Pred. No. 5.8e-38;
Matches 86; Conservative 21; Mismatches 36; Indels 0; Gaps 0;
QY 12 DGPVQGIINFEQKESNGPVKVGSIKGLTEGLGHGFHVEFGDNTAGCTSGAPHNPLSRK 71
DB 2 NGDAKGTVPFQECAGAPVKVCGEITGLSKQGHGFHVEFGDNTGCMSSGPHNPLNKE 61

QY 72 HGGPKDEERHVGDLGNWTADKDGADVSIEDSVISLSDHCIIIGRTLTVVHERADDLKGK 131
Db 62 HGAPADENRHLGLDGNIEAPGDPKVCINDCKITLFEHSIVGRTVVVHADPDDLKGK 121
QY 132 NEESTKTNAGSRLACGVIGIAQ 154
Db 122 HELSKSTGNAGARIGCGVIGICK 144

Search completed: January 28, 2003, 14:18:27
Job time : 8.78203 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 13:58:08 ; Search time 8.58126 Seconds
(without alignments)
2391.326 Million cell updates

```

Title: US-09-904-987-6
Perfect score: 821
Sequence: 1 MATKAVCVLKGDPVQGIIN.....STKTGNAGSRACGVIGIAQ 154

```

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 sub

```
Database :
1: /SDS2/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*
2: /SDS2/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*
3: /SDS2/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*
4: /SDS2/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*
5: /SDS2/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*
6: /SDS2/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*
7: /SDS2/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*
8: /SDS2/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*
9: /SDS2/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*
10: /SDS2/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*
11: /SDS2/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*
12: /SDS2/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*
13: /SDS2/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*
14: /SDS2/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*
15: /SDS2/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*
16: /SDS2/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*
17: /SDS2/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*
18: /SDS2/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*
19: /SDS2/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*
20: /SDS2/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
21: /SDS2/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*
22: /SDS2/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*
23: /SDS2/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARY

Result No.	Score	% Match	Query Length	DB Length	ID	Description
1	821	100.0	154	8	AAFP07929	Sequence of human
2	821	100.0	154	15	AAAE1336	Human superoxide d
3	821	100.0	154	17	AAAF91931	Human superoxide d
4	821	100.0	154	20	AAAG96318	Cu/Zn superoxide d
5	821	100.0	154	23	AAU10377	Superoxide dismuta
6	821	100.0	160	11	AAAF06608	Human Superoxide D
7	821	100.0	429	19	AAAF33617	SOD/p31pol structu
8	821	100.0	429	21	AAAF77311	HIV-1 p31 pol/huma
9	821	100.0	618	19	AAAW48909	SOD-1/TTC hybrid p
10	821	100.0	1021	18	AAW34481	HCV antigen combin

11	821	100.0	1021	19	AAM40039	Fusion protein c20
12	821	100.0	1021	23	AAE22050	pSOD/c200/core exp
13	821	100.0	1099	23	AAP76378	HCV multiple epitope
14	818	99.6	154	9	AAP61019	Sequence of human
15	816	99.4	153	7	AAP80235	Sequence encoded b
16	816	99.4	153	7	AAP60784	Human super-oxide
17	816	99.4	153	8	AAP71249	Human superoxide d
18	816	99.4	153	12	AAR11811	Recombinant human
19	816	99.4	153	15	AAR47062	Human superoxide d
20	816	99.4	153	18	AAW22113	Human Cu/ZnSOD. H
21	816	99.4	153	19	AAW47127	Human cytoplasmic
22	816	99.4	153	20	AAW89197	Superoxide dismuta
23	816	99.4	153	22	AAB46775	Human superoxide-d
24	816	99.4	153	23	AAM52492	Superoxide dismuta
25	816	99.4	174	9	AAP81018	Sequence of signal
26	816	99.4	275	19	AAS33615	SOD/env-5b protein
27	816	99.4	275	21	AAW77309	HIV-1 env5b/human
28	816	99.4	392	19	AAW33618	SOD/env-4 fusion p
29	815	99.3	154	14	AAW17901	Human superoxide d
30	813	99.0	153	5	AAP40223	Human cytoplasmic
31	813	99.0	154	11	AAR05283	Amino acid sequenc
32	812	98.9	154	9	AAP81266	Human Cu, Zn, supe
33	812	98.9	160	13	AAW24100	HB-SOD-4. Homo sa
34	812	98.9	171	13	AAW26510	HB-SOD derivative
35	810	98.7	161	13	AAW29645	SOD. Synthetic.
36	809	98.5	392	21	AAW77312	HIV-1 env4/human s
37	807	98.3	171	11	AAW07119	SOD analogue with
38	806	98.2	153	15	AAW51388	Human SOD. Homo s
39	806	98.2	180	11	AAW07117	SOD analogue with
40	806	98.2	180	11	AAW07118	SOD analogue with
41	804	97.9	180	13	AAW26509	HB-SOD derivative
42	801.5	97.6	152	14	AAW34899	Cytoplasmic supero
43	801	97.6	153	11	AAW06022	Superoxidisedismut
44	797	97.1	183	13	AAW27948	GAG fusion protein
45	797	97.1	183	13	AAW27951	GAG fusion protein

ALIGNMENTS

RESULT 1	
AAP70929	ID AAP70929 standard; Protein; 154 AA.
XX	AC
XX	AAAP70929;
XX	03-MAY-1991 (first entry)
XX	Sequence of human superoxidase dismutase (hsOD).
DE	Enzyme; arthro-rheumatoid osteoarthritis; radiation-induced effects.
XX	Homo sapiens.
XX	OS
PN	DE3628508-A.
XX	PD
XX	12-MAR-1987.
XX	22-AUG-1986; 86DE-3628508.
XX	23-AUG-1985; 85JP-0185246.
PR	15-AUG-1986; 86JP-0191235.
XX	(TOXN) TOYO JOZO KK.
XX	Sagai H, Takahara M, Katsuragi S, Kajiwara J, Masujima H;
XX	WPI; 1987-073705/11.
DR	N-PSDB; AAN71204.
XX	New polypeptide analogues of human superoxidisedismutase - and metal cong. dimers, useful e.g. for treating osteoarthritis
XX	

PS Disclosure; Figure 3; 25pp; German.
 XX
 CC cDNA encoding hSOD (AA71204) is isolated using a 5-triplet probe
 CC (AA71205). The cDNA is subjected to site-specific mutagenesis and
 CC expressed to produce hSOD analogues (AAP70930) which are claimed.
 XX
 SQ Sequence 154 AA;
 Query Match 100.0%; Score 821; DB 8; Length 154;
 Best Local Similarity 100.0%; Pred. No. 2.1e-86;
 Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVEFGDNTAGCTS 60
 DB 1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVEFGDNTAGCTS 60
 QY 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVLSGDHCHIIIGRTLTV 120
 DB 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVLSGDHCHIIIGRTLTV 120
 QY 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
 DB 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154

RESULT 2
 AAR61336
 ID AAR61336 standard; Protein; 154 AA.
 XX
 AC AAR61336;
 XX
 DT 31-MAY-1995 (first entry)
 XX
 DE Human superoxide dismutase SOD-1.
 XX
 KW Human superoxide dismutase; hSOD1; neurodegeneration;
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease;
 KW Hallervorden-Spatz disease; olivopontocerebellar atrophy;
 KW familial amyotrophic lateral sclerosis; FALS; diagnosis; mutant SOD;
 KW reverse transcription polymerase chain reaction; SSCP analysis.
 XX
 OS Homo sapiens.
 XX
 PN W09419493-A.
 XX
 PD 01-SEP-1994.
 XX
 PF 28-FEB-1994; 94WO-US02089.
 XX
 PR 26-FEB-1993; 93US-0023980.
 XX
 PA (GEHO) GEN HOSPITAL CORP.
 XX
 PA (MASI) MASSACHUSETTS INST TECHNOLOGY.
 XX
 PI Brown R, Horvitz HR, Rosen DR;
 XX
 DR WPI; 1994-294353/36.
 DR N-PSDB; AAQ67473.
 XX
 XX
 XX Diagnosis, treatment and prevention of diseases of cell death -
 PT e.g. amyotrophic lateral sclerosis, which are the result of e.g.
 PT decreased SOD activity
 XX
 PS Claim 22; ; 94pp; English.
 XX
 CC The presence of a mutation in a gene encoding a superoxide
 CC dismutase (SOD1, SOD2 or SOD3) indicates an increased likelihood of
 CC developing a cell death disease, specifically a neurodegenerative
 CC disease. The use of SOD polypeptides to treat amyotrophic lateral
 CC sclerosis and diseases involving a deleterious mutation in the
 CC glutathione peroxidase-encoding gene, the catalase-encoding gene
 CC and the nitric oxide-encoding gene is claimed. The specification
 CC includes the sequences of human SOD1, hSOD2 and hSOD3 (AAR61336-
 CC

CC AAR61338, respectively).
 XX
 SQ Sequence 154 AA;
 Query Match 100.0%; Score 821; DB 15; Length 154;
 Best Local Similarity 100.0%; Pred. No. 2.1e-86;
 Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVEFGDNTAGCTS 60
 DB 1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVEFGDNTAGCTS 60
 QY 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVLSGDHCHIIIGRTLTV 120
 DB 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVLSGDHCHIIIGRTLTV 120
 QY 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
 DB 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154

RESULT 3
 AAR91931
 ID AAR91931 standard; Protein; 154 AA.
 XX
 AC AAR91931;
 XX
 DT 16-DEC-1996 (first entry)
 XX
 DE Human superoxide dismutase.
 XX
 KW Human; superoxide dismutase; hSOD; Saccharomyces cerevisiae;
 KW molasses; medium; recombinant production; microbiology;
 KW medicine.
 KW
 XX
 OS Homo sapiens.
 XX
 PN RU2044771-C1.
 XX
 PD 27-SEP-1995.
 XX
 PF 01-JUN-1993; 93RU-0028609.
 XX
 PR 01-JUN-1993; 93RU-0028609.
 XX
 PA (TRIS-) TRIS STOCK CO LTD.
 XX
 PI Efremov VD, Kozlov DG, Vinetskii YP;
 XX
 DR WPI; 1996-228796/23.
 DR N-PSDB; AAT29715.
 XX
 PT Human superoxide dismutase-producing Saccharomyces cerevisiae
 PT strain - can grow on medium which contains molasses and produces
 PT hSOD yields corresp. to 20 % of cell protein
 XX
 PS Disclosure; Columns 5-6; 6pp; Russian.
 XX
 CC The present sequence is human superoxide dismutase (hSOD), the cDNA
 CC of which was used in the prepn. of a Saccharomyces cerevisiae
 CC strain which can grow on a medium contg. molasses, and produces
 CC hSOD yields corresp. to 20 % of cell protein. The recombinant hSOD
 CC produced can be used in the microbiological, and medical
 CC industries.
 XX
 SQ Sequence 154 AA;
 Query Match 100.0%; Score 821; DB 17; Length 154;
 Best Local Similarity 100.0%; Pred. No. 2.1e-86;
 Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVEFGDNTAGCTS 60
 DB 1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVEFGDNTAGCTS 60

Db ' 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVFEGDNTAGCTS 60
 QY 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSTEDSVISLSDGHCIIIGRTLTV 120
 Db 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSTEDSVISLSDGHCIIIGRTLTV 120
 QY 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
 Db 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154

RESULT 4
 AAW96318
 ID AAW96318 standard; Protein; 154 AA.
 XX
 AC AAW96318;
 XX
 DT 28-JUN-1999 (first entry)
 DE
 XX Cu/Zn superoxide dismutase.
 XX Manganese containing superoxide dismutase; MnSOD; IDDM;
 KW diabetes mellitus; treatment; nitric oxide; NO; beta cell;
 KW fatty acid; lipotoxic; cytotoxic; cytokine; osteoporosis;
 KW inflammatory disease; autoimmune disease; neurodegenerative disease.
 XX Homo sapiens.
 OS
 XX WQ9906059-A2.
 PN
 XX 11-FEB-1999.
 PD
 XX 30-JUL-1998; 98WO-US15781.
 XX
 PF
 XX 03-MAR-1998; 98US-0055092.
 PR
 XX 30-JUL-1997; 97US-0055092.
 XX
 XX (BETA-) BETAGENE INC.
 PA (TEXA) UNIV TEXAS SYSTEM.
 XX
 XX Clark SA, Hohmeier H, Koyama K, Lee Y, Newgard CB;
 PI Ohneda M, Shmabukurom, Thigpen A, Unger RH;
 PI
 XX WPI: 1999-153448/13.
 DR N-PSDB; AAX08430.
 XX
 XX Protection of mammalian cells against immunotoxicity or lipotoxicity
 PT - used for treating, e.g. diabetes, obesity, wasting syndromes,
 PT osteoporosis, inflammatory diseases, autoimmune diseases or
 PT neurodegenerative diseases
 XX
 PS Disclosure; Page 235-236; 253pp; English.
 XX
 CC Inhibition of cytokine mediated immunotoxicity of cells can be
 CC achieved by blocking free radical production or the accumulation of
 CC free radicals in that cell. Treatment of insulin dependent diabetes
 CC mellitus (IDDM) can be achieved by blocking nitric oxide (NO)
 CC production in a pancreatic beta cell and by providing a composition
 CC comprising an agent that reduces levels of fatty acids in the cells
 CC and protects beta-cells of the subject against lipid-mediated cell
 CC death. Cells can also be protected against nitric oxide mediated
 CC cytotoxicity by introducing into the cell an antioxidant agent.
 CC The methods can be used for protecting cells against immunotoxicity
 CC mediated by, e.g. IL-1 beta, IL-1 alpha, gamma IFN, TNF alpha, TNF
 CC beta, IL-8, IL-2, IL-6, IL-3, IL-5, IL-7, IL-9, IL-14, IL-17,
 CC granulocyte-macrophage colony stimulating factor or monocyte
 CC chemoattractant protein-1. The methods can be used for the treatment
 CC of e.g. insulin-dependent diabetes mellitus (IDDM), NIDDM, obesity,
 CC wasting syndromes, short stature, osteoporosis, inflammatory
 CC diseases, autoimmune diseases, or neurodegenerative diseases.
 XX
 SQ Sequence 154 AA;

Query Match 100.0%; Score 821; DB 20; Length 154;
 Best Local Similarity 100.0%; Pred. No. 21e-86;
 Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVFEGDNTAGCTS 60
 Db 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVFEGDNTAGCTS 60
 QY 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSTEDSVISLSDGHCIIIGRTLTV 120
 Db 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSTEDSVISLSDGHCIIIGRTLTV 120
 QY 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
 Db 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154

RESULT 5
 AAW10377
 ID AAW10377 standard; Protein; 154 AA.
 XX
 AC AAW10377;
 XX
 DT 14-FEB-2002 (first entry)
 XX
 DE Superoxide dismutase 1 (SOD1).
 XX
 KW Superoxide dismutase 1; soluble amyotrophic lateral sclerosis 1 (adult);
 KW haplotyping; SOD1.
 XX Homo sapiens.
 OS
 XX WO200185741-A2.
 PN
 XX 15-NOV-2001.
 PD
 XX 07-MAY-2001; 2001WO-US14772.
 PF
 XX 05-MAY-2000; 2000US-202491P.
 PR
 XX (GENA-) GENAISSANCE PHARM INC.
 PA
 XX Choi JY, Bentivegna SC, Klien SE, Koshy B, Parks KE;
 PI
 XX WPI: 2002-055578/07.
 DR N-PSDB; AAS16687.
 XX
 XX Isolated human superoxide dismutase 1 (SOD1) soluble polynucleotide,
 PT useful for screening therapeutic compounds, comprises a sequence which
 PT is a polymorphic variant of reference sequence for the SOD1 gene or its
 PT fragment -
 XX
 PS Disclosure; Fig 3; 70pp; English.
 XX
 CC The invention relates to an isolated human superoxide dismutase 1,
 CC soluble (amyotrophic lateral sclerosis 1 (adult)) (SOD1) polynucleotide
 CC (1) comprising a sequence which is a polymorphic variant of a reference
 CC sequence for the SOD1 gene. Haplotyping the SOD1 gene of an individual,
 CC involves: (a) determining whether the individual has one of the SOD1
 CC haplotypes or haplotype pairs given in the specification; or
 CC (b) determining for one copy of the SOD1 gene present in the individual,
 CC the identity of the nucleotide at two or more polymorphic sites selected
 CC from PSI-7. The method is useful for determining whether an individual
 CC has a haplotype or haplotype pairs defined in the specification. The
 CC method is also useful for improving the efficacy and reliability of
 CC several steps in the discovery and development of drugs for treating
 CC diseases associated with SOD1 activity, e.g., amyotrophic lateral
 CC sclerosis, and to validate SOD1 as a candidate agent for treating a
 CC specific condition or disease associated with SOD1 activity. It can
 CC further be used in the design of clinical trials of candidate drugs for
 CC treating a specific condition or disease predicted to be associated with
 CC SOD1 activity. (1) is useful in studying the expression and function of
 CC SOD1, and in expressing SOD1 protein for use in screening for candidate

CC This is a SOD/p31pol structural region contained in plasmid pCl/1-pSP31-
 CC ADH-GAP. This is used in the construction of a vector for expression of a
 CC polypeptide in a mammalian cell, comprising a polypeptide coding sequence
 CC operably linked downstream of an enhanced promoter. The enhanced promoter
 CC comprises the human cytomegalovirus immediate early region (HCMV IE1)
 CC promoter and the first intron proximate to the 3' end of the HCMV IE1
 CC promoter. The polypeptide can be any of the HIV recombinant polypeptides
 CC and especially HIV gp120. Expression of HIV gp120 by COS 7 cells
 CC transfected with pCMV6a containing the gp120 coding region, where pCMV6a
 CC is a vector containing the above enhanced promoter, is increased by a
 CC factor of 50-100 compared with the use of a vector containing the SV40
 CC early promoter.

XX Sequence 429 AA;

Query Match 100.0%; Score 821; DB 19; Length 429;
 Best Local Similarity 100.0%; Pred. No. 8.6e-86;
 Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVHFGDNTAGCTS 60
 |||||
 Db 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVHFGDNTAGCTS 60
 QY 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVLSGDHCHCIIGRTLTV 120
 |||||
 Db 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVLSGDHCHCIIGRTLTV 120
 QY 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
 |||||
 Db 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154

RESULT 8

AAV77311
 ID AAY77311 standard; Protein; 429 AA.

AC AAY77311;

XX 22-MAY-2000 (first entry)

XX HIV-1 p31 pol/human SOD fusion protein encoded by pCl/1-pSP31-ADH-GAP.

XX HIV-1; immunoassay; antigen; pol fragment; p31; recombinant; antibody;
 KW detection; p31 pol; human; superoxide dismutase; SOD; fusion gene.

XX Chimeric - Human immunodeficiency virus type 1 'ATCC CRL 8597'.

OS Chimeric - Homo sapiens.

XX US6013432-A.

XX 11-JAN-2000.

XX 17-MAY-1995; 95US-0443434.

XX 08-JUL-1993; 93US-0089407.

PR 24-DEC-1987; 87US-0138894.

PR 17-AUG-1992; 92US-0931154.

PR 31-OCT-1984; 84US-0667501.

PR 30-JAN-1985; 85US-0695534.

PR 06-SEP-1985; 85US-0773447.

XX (CHIR) CHIRON CORP.

XX Luciw PA, Dina D;

XX WPI; 2000-170256/15.

DR N-PSDB; AAZ20910.

XX Immunoassay for antibodies against human immune deficiency virus, for
 PT diagnosing infection, uses an immunogenic fragment of the pol protein
 PT as antigen -

XX Example 4; Fig 24A-B; 99pp; English.

XX The invention relates to the improvement of HIV-1 immunoassays by the
 CC use of an HIV-1 antigen comprising an immunogenic fragment of
 CC recombinant or synthetic HIV-1 pol, which is encoded by an approximately
 CC 9.7 kb sequence between a BstXI restriction site at position 3006 and an
 CC NdeI site at position 5131 of the genome (the proviral DNA sequence is
 CC given in AAZ90201). The immunogenic pol fragment is not immunologically
 CC cross-reactive with human T cell lymphotropic viruses I or II. The
 CC invention also encompasses the use of p31 as an antigen. The recombinant
 CC antigens may be produced in Escherichia coli, Saccharomyces cerevisiae or
 CC in mammalian cells. Immunoassays using the recombinant HIV proteins may
 CC be used to diagnose and stage HIV-1 infections. Sequences
 CC AAZ90203-290211 represent HIV-1 (ATCC CRL 8597) genes, or fusions
 CC thereof, used in exemplifications of the present invention to construct
 CC E. coli and/or S. cerevisiae expression vectors for recombinant protein
 CC expression.

SQ Sequence 429 AA;

Query Match 100.0%; Score 821; DB 21; Length 429;
 Best Local Similarity 100.0%; Pred. No. 8.6e-86;
 Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVHFGDNTAGCTS 60
 |||||
 Db 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVHFGDNTAGCTS 60
 QY 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVLSGDHCHCIIGRTLTV 120
 |||||
 Db 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVLSGDHCHCIIGRTLTV 120
 QY 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
 |||||
 Db 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154

RESULT 9

AAW48909

ID AAW48909 standard; Protein; 618 AA.

XX AAW48909;

XX 23-SEP-1998 (first entry)

XX SOD-1/TTC hybrid protein.

XX Chimeric; copper-zinc superoxide dismutase; SOD-1; TTC; SOD:Tet451;
 KW tetanus toxin fragment C; tetanus holotoxin; nerve cell; stroke;
 KW neurological disorder; oxidative stress; brain hypoxia-reperfusion;
 KW epilepsy; Parkinson's disease; Huntington's disease.

OS Chimeric - Homo sapiens.

OS Chimeric - Clostridium tetani.

XX Key Location/Qualifiers

XX Region 1..163

XX Region /note= "SOD-1"

XX Region 168..618

XX Region /note= "TTC moiety"

XX US5780024-A.

XX 14-JUL-1998.

XX 21-JUN-1996; 96US-0668381.

XX 23-JUN-1995; 95US-0000473.

XX 21-JUN-1996; 96US-0668381.

XX (GEHO) GEN HOSPITAL CORP.

XX (UYMA-) UNIV MARYLAND BALTIMORE.

XX Brown RH, Fishman PS, Francis JW, Hosler BA;

```

XX WPI: 1998-412999/35.
DR N-PSDB; AAV32580.
XX
PT New hybrid protein of superoxide dismutase and tetanus toxin
PT fragment C - having increased uptake by neurons and retention of
PT enzymatic activity in these cells, for treating neurological
PT diseases associated with oxidative stress
XX
XX Claim 7; Columns 23-26; 23pp; English.
XX
CC The present sequence represents an enzymatically active human
CC copper-zinc superoxide dismutase (SOD-1) fused at its carboxyl
CC terminus with the tetanus toxin fragment C (TTC) moiety. The TTC
CC moiety constitutes amino acid residues 865-1315 of the tetanus
CC holotoxin. The hybrid protein, referred as SOD:Tet451, is claimed
CC to have the following properties: (a) it exhibits Cu/Zn SOD enzymatic
CC activity; (b) the TTC moiety selectively binds to nerve cells and
CC allows uptake of the hybrid protein into these cells; and (c) it
CC retains substantial SOD enzymatic activity following cellular uptake.
CC SOD:Tet451 is claimed to be useful for treating neurological disorders
CC associated with oxidative stress, e.g. stroke, brain hypoxia-reperfusion,
CC epilepsy, Parkinson's and Huntington's diseases.
XX
SQ Sequence 618 AA;
Query Match 100.0%; Score 821; DB 19; Length 618;
Best Local Similarity 100.0%; Pred. No. 1.4e-85;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATKAVCVLKGDGPVQGIINFEQKESNGPKVYVWGSIKGLTEGLHGFHVEFGDNTAGCTS 60
DB 1 MATKAVCVLKGDGPVQGIINFEQKESNGPKVYVWGSIKGLTEGLHGFHVEFGDNTAGCTS 60
QY 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 120
DB 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 120
QY 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
DB 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
QY 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
DB 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
RESULT 10
AAW34481
ID AAW34481 standard; Protein; 1021 AA.
XX
AC AAW34481;
XX
XX 16-MAR-1998 (first entry)
XX HCV antigen combination pSOD/c200/core.
XX
XX PCR primer; amplify; HCV; hepatitis c virus; antigen combination; NS3;
XX C domain; S domain; NS5; HCV polyprotein; anti-HCV antibody; detection;
XX NS4.
XX
XX Hepatitis C virus.
XX Synthetic.
XX
XX Key Location/Qualifiers
FH Misc-difference 1..154
FT /note= "hsod fragment"
FT Misc-difference 155..159
FT /note= "linker"
FT Misc-difference 160..899
FT /note= "c200 (amino acids 1192-1931 of HCV polyprotein)"
FT Misc-difference 1..902
FT /note= "linker"
FT Misc-difference 903..1021
FT /note= "c22 (amino acids 2-120 of HCV polyprotein)"
XX
XX US5683864-A.
PN

```

```

XX 04-NOV-1997.
PD XX
XX 07-JUL-1992; 92US-0910760.
PF XX
XX 07-JUL-1992; 92US-0910760.
PR XX
PR 18-NOV-1987; 87US-0122714.
PR 30-DEC-1987; 87US-0139886.
PR 26-FEB-1988; 88US-0161072.
PR 06-MAY-1988; 88US-0191263.
PR 26-OCT-1988; 88US-0263584.
PR 14-NOV-1988; 88US-0271450.
PR 17-MAR-1989; 89US-0325338.
PR 20-APR-1989; 89US-0341334.
PR 21-APR-1989; 89US-0353896.
PR 04-APR-1990; 90US-0504352.
XX (CHIR ) CHIRON CORP.
PA
XX Choo Q, Houghton M, Kuo G;
PI
XX WPI: 1997-548976/50.
XX N-PSDB; AAT99982.
DR
XX Combination of three hepatitis C virus antigens - used for detection
XX of specific antibodies to diagnose infection
XX
XX Example 6; Column 59-68; 57pp; English.
XX
CC This sequence represents a Hepatitis c virus (HCV) antigen combination of
CC the invention. The HCV antigen combination comprises an antigen (Ag1)
CC comprising the C domain (i.e. amino acids (aa) 1-120 of the HCV
CC polyprotein), or its immunologically reactive fragment containing at
CC least 8 aa. It also comprises two additional antigens from two different
CC polypeptide domains, including at least 8 aa from the NS3, NS4, S or NS5
CC domains of the polyprotein, corresponding, respectively, to aa 1050-1640;
CC 1640-2000; 120-400 and 2000-3011 of the HCV polyprotein. Alternatively,
CC Ag1 contains at least 8 aa from the 1-122 or 9-177 aa regions of the HCV
CC polyprotein. These antigen combinations are used diagnostically to detect
CC anti-HCV antibodies, using any standard immunoassay format. These antigen
CC combinations have a broader range of reactivity with antibodies than any
CC antigen individually.
XX
SQ Sequence 1021 AA;
Query Match 100.0%; Score 821; DB 18; Length 1021;
Best Local Similarity 100.0%; Pred. No. 2.9e-85;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATKAVCVLKGDGPVQGIINFEQKESNGPKVYVWGSIKGLTEGLHGFHVEFGDNTAGCTS 60
DB 1 MATKAVCVLKGDGPVQGIINFEQKESNGPKVYVWGSIKGLTEGLHGFHVEFGDNTAGCTS 60
QY 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 120
DB 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 120
QY 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
DB 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
RESULT 11
AAW40039
ID AAW40039 standard; Protein; 1021 AA.
XX
AC AAW40039;
XX
XX 26-MAY-1998 (first entry)
XX Fusion protein c200/c22.
XX
XX Hepatitis C virus C domain; HCV; immunological activity; c200/c22;
KW

```

KW NS3 domain; NS4 domain; S domain; NS5 domain; fusion protein.
 XX Synthetic.
 OS Hepatitis virus.
 XX
 PN US5712087-A.
 XX
 XX 27-JAN-1998.
 XX
 XX 12-MAY-1995; 95US-0440519.
 XX
 XX 07-JUL-1992; 92US-0910760.
 PR 04-APR-1990; 90US-0504352.
 PR 12-MAY-1995; 95US-0440519.
 XX
 XX (CHIR) CHIRON CORP.
 XX
 XX Choo Q, Houghton M, Kuo G;
 PI WPI; 1998-119973/11.
 DR N-PSDB; AAV09990.
 XX
 XX Immunoassays for hepatitis C virus antibodies - using combinations
 PT of antigenic fragments of HCV polyprotein
 XX
 XX Example 6; Fig 4; 59pp; English.
 XX
 CC This sequence represents a fusion protein constructed from the hepatitis
 CC C virus core domain (which is situated at the carboxy terminus of the
 CC fusion protein) and a c200 construct (a fusion of the NS3 and NS5
 CC domains). This protein used in the construction of novel combinations of
 CC HCV antigens that have a broader range of immunological activity than
 CC any single HCV antigen. An example of such an antigen given in this
 CC specification comprises a first antigen containing at least 8 amino acids
 CC of the C domain of the HCV polyprotein and a second antigen comprising
 CC at least 8 amino acids of the NS3 domain, the NS4 domain, the S domain or
 CC the NS5 domain of the HCV polyprotein in the form of a fusion protein, a
 CC physical mixture or bound to a solid matrix.
 XX
 SQ Sequence 1021 AA;

Query Match 100.0%; Score 821; DB 19; Length 1021;
 Best Local Similarity 100.0%; Pred. No. 2.9e-85;
 Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVEFGDNTAGCTS 60
 DB 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVEFGDNTAGCTS 60
 QY 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSGDHCIIIGRTLTV 120
 DB 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSGDHCIIIGRTLTV 120
 QY 121 HEKADDLGKGGNEESTKTGNAGSLACGVIGIAQ 154
 DB 121 HEKADDLGKGGNEESTKTGNAGSLACGVIGIAQ 154

RESULT 12
 AAE22050
 ID AAE22050 standard; Protein; 1021 AA.
 XX
 AC AAE22050;
 XX
 XX 16-JUL-2002 (first entry)
 XX
 XX pSOD/c200/core expression plasmid protein.
 XX Hepatitis C virus; HCV; antigen; C domain; polyprotein; NS3 domain;
 KW NS4 domain; S domain; NS5 domain; pSOD/c200/core plasmid.
 XX
 XX Chimeric - Hepatitis C virus.
 OS Chimeric - Unidentified.

XX Key Location/Qualifiers
 FH Region 1..154
 FT /note= "hsOD"
 FT 155..159
 FT /note= "Linker region"
 FT 160..899
 FT /note= "HCV c200"
 FT 900..902
 FT /note= "Linker region"
 FT 903..1021
 FT /note= "HCV c22"
 XX

PN US6312889-B1.

XX 06-NOV-2001.

XX 12-MAY-1995; 95US-0440549.

XX 07-JUL-1992; 92US-0910760.

PR 04-APR-1990; 90US-0504352.

XX (CHIR) CHIRON CORP.

XX Houghton M, Choo Q, Kuo G;

XX WPI; 2002-040268/05.

DR N-PSDB; AAD35044.

XX Combination of hepatitis C viral (HCV) antigens, useful in improved
 PT immunoassay for detecting HCV antibodies -

XX Example 6; Fig 4; 58pp; English.

XX The invention relates to combination of hepatitis C viral (HCV) antigens
 CC that have a broader range of immunological reactivity than any single HCV
 CC antigen. The combinations consist of an antigen from the C domain of the
 CC HCV polyprotein, and at least one additional HCV antigen from either the
 CC NS3 domain, the NS4 domain, the S domain, or the NS5 domain and are in
 CC the form of fusion protein, a simple physical mixture, or the individual
 CC antigens commonly bound to a solid matrix. The combinations of antigens
 CC provides broad range immunoassays for anti-HCV antibodies. The invention
 CC therefore provides a method for detecting antibodies to HCV in a mammal
 CC suspected of containing such antibodies. The present sequence is a
 CC protein encoded by pSOD/c200/core expression plasmid DNA containing HCV
 CC coding sequence.

XX Sequence 1021 AA;

Query Match 100.0%; Score 821; DB 23; Length 1021;
 Best Local Similarity 100.0%; Pred. No. 2.9e-85;
 Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVEFGDNTAGCTS 60
 DB 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVEFGDNTAGCTS 60

QY 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSGDHCIIIGRTLTV 120
 DB 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSGDHCIIIGRTLTV 120

QY 121 HEKADDLGKGGNEESTKTGNAGSLACGVIGIAQ 154
 DB 121 HEKADDLGKGGNEESTKTGNAGSLACGVIGIAQ 154

RESULT 13
 AAU76378

ID AAU76378 standard; Protein; 1099 AA.

XX AAU76378;

XX 08-MAY-2002 (first entry)

XX HCV multiple epitope fusion antigen (MEFA) 7.1 protein sequence.
 DE Hepatitis C virus; HCV; NS3/4a conformational epitope; seroconversion;
 KW Immunoassay solid support; multiple epitope fusion antigen; MEFA;
 KW non-structural protein.

XX Hepatitis C virus.
 OS Synthetic.

XX W0200196870-A2.

XX 20-DEC-2001.

XX 14-JUN-2001; 2001WO-0519156.

XX 15-JUN-2000; 2000US-212082P.

PR 02-APR-2001; 2001US-280811P.

PR 02-APR-2001; 2001US-280867P.

XX (CHIR) CHIRON CORP.

XX Chien DY, Arcangel P, Tandeske L, George-nascimento C, Coit D;
 PI Medina-selby A;

XX WPI; 2002-090228/12.

XX N-PSDB; ABK15345.

XX Immunoassay solid support, useful for detecting hepatitis C virus
 PT infection in biological sample, comprises HCV NS3/4a conformational
 PT epitope and multiple epitope fusion antigen bound to the support -
 PT Claim 5; Fig 5; 92pp; English.

XX The present invention relates to a new immunoassay solid support
 CC consisting essentially of at least one hepatitis C virus (HCV) NS3/4a
 CC conformational epitope and a multiple epitope fusion antigen (MEFA),
 CC bound to the support. The NS3/4a conformational epitope and/or
 CC MEFA reacts specifically with anti-HCV antibodies present in a biological
 CC sample from an HCV-infected individual. The immunoassay of the invention
 CC is useful for detecting hepatitis C virus infection in a biological
 CC sample. The method of the invention provides a sensitive, accurate
 CC diagnostic and prognostic tool to provide adequate patient care and to
 CC prevent transmission of HCV by blood and by blood products, or by
 CC personal contact. Use of NS3/4a conformational epitope in combination
 CC with MEFA, provides a sensitive and reliable method for detecting early
 CC HCV seroconversion. Use of MEFA has the added advantages of decreasing
 CC masking problems, improving sensitivity in detecting antibodies by
 CC allowing a greater number of epitopes on a unit surface area of
 CC substrate, and improving substrate. Detection accuracy is increased and
 CC the incidence of false results is reduced because of the identification
 CC and the use of highly immunogenic HCV antigens which are present during
 CC the early stages of HCV seroconversion. The present amino acid sequence
 CC represents the multiple epitope fusion antigen (MEFA) 7.1 of the
 CC invention.

XX Sequence 1099 AA;

Query Match 100.0%; Score 821; DB 23; Length 1099;
 Best Local Similarity 100.0%; Pred. No. 3.2e-85;
 Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATKAVCVLKGDPVQGIINFEQKESGPKVKGSIKGLTEGLHGFHVEFGDNTAGCTS 60
 |||
 Db 1 MATKAVCVLKGDPVQGIINFEQKESGPKVKGSIKGLTEGLHGFHVEFGDNTAGCTS 60

QY 61 AGPHFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDHCHIIIGRTLW 120
 |||
 Db 61 AGPHFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDHCHIIIGRTLW 120

QY 121⁶ HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
 |||
 Db 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154

RESULT 14

AAP81019

ID AAP81019 standard; protein; 154 AA.

XX AAP81019;

XX 01-NOV-1990 (first entry)

XX Sequence of human superoxidisedismutase (hsOD) encoded by cDNA.

XX Human superoxidisedismutase.

XX Key Location/Qualifiers
 FT Region 2..154
 FT /note="Claimed in claim 3"

XX JPC3237790-A.

XX 04-OCT-1988.

XX 27-MAR-1987; 87JP-0073180.

XX 27-MAR-1987; 87JP-0073180.

XX (TOXN) TOYO JOZO KK.

XX WPI; 1988-320034/45.

XX N-PSDB; AAN81345.

XX Prodn. of extracellular protein prod. -
 PT using human superoxidisedismutase linked to OMPA signal peptide
 PT and gram-negative bacteria

XX ; Fig 3 Page 508; 23pp; Japanese.

XX A gene which comprises the DNA encoding the signal peptide of the OMPA
 CC (AAN81344) linked upstream of DNA encoding an intracellular protein or
 CC peptide, esp. human superoxidisedismutase (hsOD) (AAN81345) is new. Prodn.
 CC of the intracellular protein/peptide comprises culturing bacteria contg.
 CC the new gene and accumulating the peptide/protein outside the cytoplasm
 CC of bacteria. Pref. the DNA having the basic sequence coding hsOD or its
 CC deriv. is obtd. by cloning cDNA of hsOD derived from human liver.
 CC Advantage is that purificn. of obtd. hsOD is easily carried out and the
 CC yield of the protein is higher than using conventional method.

XX Sequence 154 AA;

Query Match 99.6%; Score 818; DB 9; Length 154;
 Best Local Similarity 99.4%; Pred. No. 4.6e-86;
 Matches 153; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATKAVCVLKGDPVQGIINFEQKESGPKVKGSIKGLTEGLHGFHVEFGDNTAGCTS 60
 |||
 Db 1 MATKAVCVLKGDPVQGIINFEQKESGPKVKGSIKGLTEGLHGFHVEFGDNTAGCTS 60

QY 61 AGPHFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDHCHIIIGRTLW 120
 |||
 Db 61 AGPHFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDHCHIIIGRTLW 120

QY 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
 |||
 Db 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154

RESULT 15

AAP60235

ID AAP60235 standard; Protein; 153 AA.

XX AAP60235;

XX 15-AUG-1991 (first entry)

Search completed: January 28, 2003, 14:14:14
Job time : 11.5813 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:19:14 ; Search time 2.01912 Seconds
(without alignments)
1539.035 Million cell updates/sec

Title: US-09-904-987-6
Perfect score: 821
Sequence: 1 MATKAVCVLKGDPVQGIIN.....STKTGNAGSLRACGVIGIAQ 154

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 122226 seqs, 20178551 residues

Total number of hits satisfying chosen parameters: 122226

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	821	100.0	154	10	US-09-904-987-6
2	821	100.0	1099	10	US-09-881-654-4
3	807	99.3	152	10	US-09-883-985-7
4	655.5	79.8	151	10	US-09-883-985-10
5	638	77.7	150	10	US-09-883-985-4
6	522	63.6	150	10	US-09-883-985-11
7	492	59.9	151	10	US-09-883-985-3
8	483	58.8	151	10	US-09-883-985-6
9	482	58.7	150	10	US-09-883-985-9
10	444.5	54.1	151	10	US-09-883-985-5
11	444.5	54.1	151	10	US-09-883-985-8
12	444	54.1	152	10	US-09-883-985-12
13	382.5	46.6	829	9	US-09-881-239-5
14	365	44.5	288	9	US-09-964-899-21
15	365	44.5	310	10	US-09-925-297-795
16	358	43.6	255	10	US-09-883-985-2
17	206.5	25.2	68	9	US-09-987-675-2
18	206.5	25.2	68	10	US-09-987-655-2
19	118	14.4	25	10	US-09-839-884-54

20	75.5	9.2	439	9	US-09-738-626-6639	Sequence 6639, Ap
21	75	9.1	14	9	US-09-826-290-262	Sequence 262, App
22	75	9.1	360	10	US-09-815-242-11254	Sequence 11254, A
23	72.5	8.8	1325	10	US-09-741-669-304	Sequence 304, App
24	72	8.8	261	10	US-09-940-037A-29	Sequence 29, Appl
25	72	8.8	552	9	US-09-961-721-2	Sequence 2, Appl
26	70.5	8.6	592	9	US-09-771-382-10	Sequence 10, Appl
27	70	8.5	579	9	US-10-108-605-215	Sequence 215, App
28	68.5	8.3	134	10	US-09-731-872-474	Sequence 474, App
29	68.5	8.3	442	9	US-10-117-417-6	Sequence 6, Appl
30	68.5	8.3	442	9	US-10-117-417-16	Sequence 16, Appl
31	68	8.3	170	12	US-10-109-885-3	Sequence 3, Appl
32	68	8.3	711	9	US-09-738-626-3507	Sequence 3507, A
33	67.5	8.2	595	10	US-09-864-761-38018	Sequence 38018, A
34	67.5	8.2	727	10	US-09-833-790-397	Sequence 397, App
35	67.5	8.2	1242	9	US-09-832-292-41	Sequence 41, Appl
36	67.5	8.2	26926	9	US-09-759-508B-2	Sequence 2, Appl
37	66.5	8.1	541	9	US-09-771-382-34	Sequence 34, Appl
38	66.5	8.1	592	9	US-09-771-382-2	Sequence 2, Appl
39	66.5	8.1	592	10	US-09-797-862-17	Sequence 17, Appl
40	66	8.0	135	10	US-09-815-242-10485	Sequence 10485, A
41	66	8.0	433	10	US-09-771-161A-229	Sequence 229, App
42	66	8.0	433	10	US-09-771-161A-230	Sequence 230, App
43	66	8.0	2568	10	US-09-866-108-3	Sequence 3, Appl
44	65.5	8.0	182	10	US-09-784-810A-29	Sequence 29, Appl
45	65.5	8.0	580	9	US-09-738-626-6264	Sequence 6264, Ap

ALIGNMENTS

RESULT 1

US-09-904-987-6

; Sequence 6, Application US/09904987

; Patent No. US20020037908A1

; GENERAL INFORMATION:

; APPLICANT: No. US20020037908A1actyl, Inc.

; TITLE OF INVENTION: Methods and Compositions for Controlling Pathological and Prop

; FILE REFERENCE: 42108/26146

; CURRENT APPLICATION NUMBER: US/09/904, 987

; CURRENT FILING DATE: 2001-07-12

; NUMBER OF SEQ ID NOS: 7

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 6

; LENGTH: 154

; TYPE: PRT

; ORGANISM: homo sapiens

; PUBLICATION INFORMATION:

; DATABASE ACCESSION NUMBER: NCBI ENTREZ / P00441

; DATABASE ENTRY DATE: 2000-05-30

; RELEVANT RESIDUES: (1)..(154)

US-09-904-987-6

Query Match	100.0%	Score 821	DB 10	Length 154
Best Local Similarity	100.0%	Pred. No. 4.3e-82		
Matches 154	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1	MATKAVCVLKGDPVQGIINFEQKESNPVKVWGSIKGLTEGLHGFHVHFEFGDNTAGCTS	60	
Db	1	MATKAVCVLKGDPVQGIINFEQKESNPVKVWGSIKGLTEGLHGFHVHFEFGDNTAGCTS	60	
QY	61	AGPHFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSTEDSVISLSGDHICIGRTLTV	120	
Db	61	AGPHFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSTEDSVISLSGDHICIGRTLTV	120	
QY	121	HEKADDLGKGNNESTKTGNAGSLRACGVIGIAQ	154	
Db	121	HEKADDLGKGNNESTKTGNAGSLRACGVIGIAQ	154	

RESULT 2

US-09-881-654-4


```

; REFERENCE/DOCKET NUMBER: 1488.1020003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 150 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
; US-09-883-985-4
;
Query Match          77.7%; Score 638; DB 10; Length 150;
Best Local Similarity 82.0%; Pred. No. 3.2e-62;
Matches 123; Conservative 9; Mismatches 16; Indels 2; Gaps 1;
;
QY   3   TKAVCVLKGDPVQGIIINFEQKESNGPVKVMGSIKGLTEGLHGPHVFHFGDNTAGCT'SAG 62
      |||||:|||||:||| | | |||||:|||||:|||||:|||||
Db    1   TKAVCVLKGDPVQGTIHFPEAK--GDTVVVTGSTTGLTEGDHGFHVHQFDNTOGCT'SAG 58
;
QY   63   PHNPILSRKHGGPKDEHRHVGDIIGNVTADKDGVADVVSIEDSVLSLDGHCLIGRTLVVHE 122
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db    59   PHNPILSRKHGGPKDEHRHVGDIIGNVTADKNGAIVDIVDPLISLSGEYSIIGRTMVVHE 118
;
QY   123  KADDLGKGGNESTKTGNAGSLACGVIGI 152
      | |||||:|||||:|||||:|||||:|||||
Db    119  KPDDLGRGGNESTKTGNAGSLACGVIGI 148
;
RESULT 6
US-09-883-985-11
; Sequence 11, Application US/098833985
; Patent No. US20020081288A1
; GENERAL INFORMATION:
; APPLICANT: YU, GUO-LIANG
; ROSEN, CRAIG A.
; FRASER, CLAIRE M.
; GOCAYNE, JEANNINE D.
; TITLE OF INVENTION: SUPEROXIDE DISMUTASE-4
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNER, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET : 1100 NEW YORK AVENUE, N.W., SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/883,985
; FILING DATE: 20-Jun-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/203,607
; FILING DATE: 02-DEC-1998
; APPLICATION NUMBER: US 08/722,050
; FILING DATE: 23-JAN-1997
; APPLICATION NUMBER: US 08/225,757
; FILING DATE: 11-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: STEFFE, ERIC K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.1020003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 11:

```

```

;
; SEQUENCE CHARACTERISTICS:
;
; LENGTH: 150 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; MOLECULE TYPE: protein
;
; SEQUENCE DESCRIPTION: SEQ ID
US-09-883-985-11

```

Query Match	63.6%	Score	522;	DB	10;	Length	150;
Best Local Similarity	66.9%;	Pred. No.	1.3e-49;				
Matches	99;	Conservative	17;	Mismatches	30;	Indels	2; Gaps
QY	4	KAVCVLKGDPVQGIINFEOKESKPKVWGSIKGLTEGLHGHHVFHFGDNTAGCTSACP	63				
		: :: :: : : ::					
Db	2	KAVCVLAGSDGVKGYRFEQQD-DGDVTVEGIEGLTGNHGHFHVFDGNTNGCLSLAP	60				
QY	64	HFNPLSRKIGGPKDEERHWGLGNVTADKCGADVSTEDSVISLSGDHCITGRTLLVWEHK	123				
Db	61	HFNPNKNHSGPKADRHRVGDLGNVTAE-CGVAQFNFTDPQISLKGERSTIGTAVVHEK	119				
QY	124	ADDLKGKGNNESTKKTGNAGSRACGVIG	151				
Db	120	QDDLGKGGDDSLKTCNAGGRACGVIG	147				

RESULT 7
US-09-883-985-3
; Sequence 3, Application US/09883985
; Patent No. US20020081288A1
; GENERAL INFORMATION:

```

; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-883-985-3

Query Match          59.9%   Score 492;   DB 10;   Length 153;
Best Local Similarity 60.0%;   Pred. No. 2.4e-46;
Matches 90; Conservative 18; Mismatches 42; Indels 0; Caps 0;

Qy  4  KAVCVLKGDPVQGIINFEQKESGPNPKVYKSIKGLTEGLHGFHVFEGDNTAGCTSAGP  63
      ||||| : || : || : || : || : || : || : || : || : || : || : || : ||
Db  2  KAVCVMTGTAGVKGWKFQETDNGPVRVHVAEFSGLKAGKHGFHVFEGDNTGCTSAGA  61
      ||||| : || : || : || : || : || : || : || : || : || : || : ||

Qy  64  HFNPLSRKHHGPKDEERHVGDLGNVTKDKGVADVSIEDSVISLSDGCHCIIGRTLVVHEK  123
      ||||| : || : || : || : || : || : || : || : || : || : || : || : ||
Db  62  HFNPTKGHEGAPEDSIRHVGDLGNVAGADGNVYNTDKLISNGSHSIIGRSMVHEN  121
      ||||| : || : || : || : || : || : || : || : || : || : || : || : ||

Qy  124  ADDLKGKGNNESTTKTNAGSLACGVIGIA  153
      ||||| : || : || : || : || : || : || : || : || : || : || : || : ||
Db  122  EDDLGRGSHLSKVITGNAGRLAGVVGLA  151
      ||||| : || : || : || : || : || : || : || : || : || : || : || : ||

```

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999

```

Query Match 58.88; Score 483; DB 10; Length 151;

Best Local Similarity 61.7%; Pred. No. 2.2e-45;
Matches 92; Conservative 18; Mismatches 37; Indels 2; Gaps 1;

Qy 4 KAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVFHFGDNTAGCTAGP 63
Db 2 KAVCVLNGD--AKGVTFQESSGTPVKVSGVEVCGGLAKGLHGFHVFHFGDNTAGCTAGP 59

Qy 64 HFNPLSRKHGPKDEERHVGDLGNVTADKDGADVSIEDSVISLSDGHCIIIGRTLVVHEK 123
Db 60 HFNYPKGHGAPVDENRHLGDLGNIEATGDCPTKYNITDSKITLFGADSIIGRTVVVHAD 119

Qy 124 ADDLGKGGNEESTKTGNAGSRACGVIGI 152
Db 120 ADDLGKGGHLSKSTGNAGRGVACGIIGI 148

RESULT 9
US-09-883-985-9
; Sequence 9, Application US/09883985
; Patent No. US20020081288A1
; GENERAL INFORMATION:
; APPLICANT: YU, GUO-LIANG
; ROSEN, CRAIG A.
; FRASER, CLAIRE M.
; GOCAYNE, JEANNINE D.
; TITLE OF INVENTION: SUPEROXIDE DISMUTASE-4
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, N.W., SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/883,985
; FILING DATE: 20-Jun-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/203,607
; FILING DATE: 02-DEC-1998
; APPLICATION NUMBER: US 08/722,050
; FILING DATE: 23-JAN-1997
; APPLICATION NUMBER: US 08/225,757
; FILING DATE: 11-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: STEFFE, ERIC K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.1020003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 150 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 9:
US-09-883-985-9

Query Match 58.7%; Score 482; DB 10; Length 150;
Best Local Similarity 63.1%; Pred. No. 2.8e-45;
Matches 94; Conservative 15; Mismatches 38; Indels 2; Gaps 2;

Qy 4 KAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVFHFGDNTAGCTAGP 63
Db 2 KAVCVLNGD--AKGVTFQESSGTPVKVSGVEVCGGLAKGLHGFHVFHFGDNTAGCTAGP 59

Qy 64 HFNPLSRKHGPKDEERHVGDLGNVTADKDGADVSIEDSVISLSDGHCIIIGRTLVVHEK 123
Db 60 HFNYPKGHGAPVDENRHLGDLGNIEATGDCPTKYNITDSKITLFGADSIIGRTVVVHAD 119

Qy 124 ADDLGKGGNEESTKTGNAGSRACGVIGI 152
Db 120 ADDLGKGGHLSKSTGNAGRGVACGIIGI 148

RESULT 10
US-09-883-985-5
; Sequence 5, Application US/09883985
; Patent No. US20020081288A1
; GENERAL INFORMATION:
; APPLICANT: YU, GUO-LIANG
; ROSEN, CRAIG A.
; FRASER, CLAIRE M.
; GOCAYNE, JEANNINE D.
; TITLE OF INVENTION: SUPEROXIDE DISMUTASE-4
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, N.W., SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/883,985
; FILING DATE: 20-Jun-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/203,607
; FILING DATE: 02-DEC-1998
; APPLICATION NUMBER: US 08/722,050
; FILING DATE: 23-JAN-1997
; APPLICATION NUMBER: US 08/225,757
; FILING DATE: 11-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: STEFFE, ERIC K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.1020003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 151 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-883-985-5

Query Match 54.1%; Score 444.5; DB 10; Length 151;
Best Local Similarity 57.0%; Pred. No. 3.4e-41;
Matches 85; Conservative 13; Mismatches 50; Indels 1; Gaps 1;

Qy 4 KAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVFHFGDNTAGCTAGP 63
Db 2 KAVCVLNSSEGVGTFTFFTH--EGNGATTVTGTVSGRLPGLHGFHVFHFGDNTAGCTAGP 60

Qy 64 HFNPLSRKHGPKDEERHVGDLGNVTADKDGADVSIEDSVISLSDGHCIIIGRTLVVHEK 123
Db 61 HFNPDGKTHCAPEDANRHHAGDLGNIIVGDGTATFTTDSQIPLSGPNSIVGRAIVVHAD 120

Search completed: January 28, 2003, 14:40:14
Job time : 3.01912 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:14:29 ; Search time 39,9197 Seconds
(without alignments)
2487.216 Million cell updates/sec

Title: US-09-904-987-6
Perfect score: 821
Sequence: 1 MATKAVCVLKGDGPVQGIIN.....STKTGNAGSLRACGVIGIAQ 154

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_Main :*

1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US097_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US099_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	821	100.0	154	1	PCT-US01-14772-3
2	821	100.0	154	1	PCT-US02-10824-198
3	821	100.0	154	1	PCT-US02-22032-6
4	821	100.0	154	21	US-09-744-016A-2
5	821	100.0	154	21	US-09-744-016A-16
6	821	100.0	154	21	US-09-791-537-38416

7	821	100.0	154	23	US-09-904-987-6	Sequence 6, Appli
8	821	100.0	156	21	US-09-760-475-2306	Sequence 2306, Ap
9	821	100.0	668	13	US-08-993-010-2	Sequence 2, Appli
10	821	100.0	839	21	US-09-791-537-91118	Sequence 91118, A
11	821	100.0	1099	22	US-09-881-654-4	Sequence 4, Appli
12	816	99.4	153	3	US-07-768-912-13	Sequence 13, Appli
13	816	99.4	153	21	US-09-744-016A-14	Sequence 14, Appli
14	816	99.4	153	24	US-10-018-770-1	Sequence 1, Appli
15	816	99.4	154	21	US-09-791-537-151133	Sequence 151133,
16	814	99.1	154	21	US-09-791-537-118562	Sequence 118562,
17	814	99.1	154	27	US-09-389-987-185	Sequence 185, App
18	814	99.1	154	27	US-09-412-418-185	Sequence 185, App
19	812	98.9	153	21	US-09-791-537-99028	Sequence 99028, A
20	808	98.4	153	21	US-09-791-537-93750	Sequence 93750, A
21	807	98.3	152	16	US-09-203-607-7	Sequence 7, Appli
22	807	98.3	152	22	US-09-883-985-7	Sequence 7, Appli
23	807	98.3	179	21	US-09-791-537-68361	Sequence 68361, A
24	802	97.7	154	21	US-09-791-537-53587	Sequence 53587, A
25	797	97.1	153	21	US-09-791-537-83751	Sequence 83751, A
26	797	97.1	154	21	US-09-791-537-152870	Sequence 152870,
27	797	97.1	171	21	US-09-791-537-50085	Sequence 50085, A
28	797	97.1	841	3	US-07-680-296-86	Sequence 86, Appli
29	797	97.1	841	22	US-09-884-455-86	Sequence 86, Appli
30	797	97.1	841	22	US-09-884-455-86	Sequence 86, Appli
31	793	96.6	153	21	US-09-791-537-38270	Sequence 38270, A
32	790	96.2	153	27	US-60-389-987-74	Sequence 74, Appli
33	790	96.2	153	27	US-60-412-418-74	Sequence 280, Appli
34	788	96.0	200	1	PCT-US01-14826-280	Sequence 13968, A
35	777	94.6	153	21	US-09-791-537-13968	Sequence 151958,
36	773	94.2	153	21	US-09-791-537-151958	Sequence 151958,
37	763	92.9	153	21	US-09-791-537-150043	Sequence 150043,
38	744	90.6	154	4	US-08-023-980-38	Sequence 38, Appli
39	744	90.6	154	6	US-08-204-052-51	Sequence 51, Appli
40	689.5	84.0	153	8	US-08-486-953-51	Sequence 51, Appli
41	684.5	83.4	152	21	US-09-791-537-42081	Sequence 42081, A
42	684.5	83.4	152	21	US-09-791-537-42336	Sequence 42336, A
43	684.5	83.4	152	21	US-09-791-537-85330	Sequence 85330, A
44	680	82.8	154	21	US-09-791-537-43187	Sequence 43187, A
45	678	82.6	154	21	US-09-791-537-43206	Sequence 43206, A

ALIGNMENTS

RESULT 1
PCT-US01-14772-3
; Sequence 3, Application PC/TUS0114772
; GENERAL INFORMATION:
; APPLICANT: Genesuisse Pharmaceuticals, Inc.
; APPLICANT: Choi, Julie Y.
; APPLICANT: Bentivegna, Steven C.
; APPLICANT: Kliem, Stefanie E.
; APPLICANT: Koshiy, Beena
; APPLICANT: Parks, Katie E.
; TITLE OF INVENTION: HAPLOTYPES OF THE SODI GENE
; FILE REFERENCE: MMH-0429PCT SODI
; CURRENT APPLICATION NUMBER: PCT/US01/14772
; PRIOR FILING DATE: 2001-05-07
; PRIOR APPLICATION NUMBER: 60/202,491
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-14772-3

Query Match 100.0%; Score 821; DB 1; Length 154;
Best Local Similarity 100.0%; Pred. No. 2.2e-84;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATKAVCVLKGDGPVQGIINFEQESNCPVKWGSIKLTGLHGFHVFHFGDNTAGCTS 60

|||||
Db 1 MATKAVCVLKGDPVQGIINFEQESNGPVKVWGSIKGLTEGLHGFHVHEFGDNTAGCTS 60
QY 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 120
Db 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 120
QY 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
Db 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154

RESULT 2

PCT-US02-10824-198
; Sequence 198, Application PC/TUS0210824
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies
; TITLE OF INVENTION: Prostate Cancer Expression Profiles
; FILE REFERENCE: 90 206 PCT
; CURRENT APPLICATION NUMBER: PCT/US02/10824
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: US 60/281,732
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: US 60/281,731
; PRIOR FILING DATE: 2001-04-06
; NUMBER OF SEQ ID NOS: 211
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 198
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-10824-198

Query Match 100.0%; Score 821; DB 1; Length 154;
Best Local Similarity 100.0%; Pred. No. 2.2e-84;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATKAVCVLKGDPVQGIINFEQESNGPVKVWGSIKGLTEGLHGFHVHEFGDNTAGCTS 60
Db 1 MATKAVCVLKGDPVQGIINFEQESNGPVKVWGSIKGLTEGLHGFHVHEFGDNTAGCTS 60
QY 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 120
Db 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 120
QY 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
Db 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154

RESULT 3

PCT-US02-22032-6
; Sequence 6, Application PC/TUS0222032
; GENERAL INFORMATION:
; APPLICANT: Novactyl, Inc.
; TITLE OF INVENTION: Methods and Compositions for Controlling Protein Assembly or Aggr
; FILE REFERENCE: 42108/34520
; CURRENT APPLICATION NUMBER: PCT/US02/22032
; PRIOR FILING DATE: 2002-07-11
; PRIOR APPLICATION NUMBER: 09/904,987
; PRIOR FILING DATE: 12 July 2001
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 154
; TYPE: PRT
; ORGANISM: homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NCBI ENTREZ / P00441
; DATABASE ENTRY DATE: 2000-05-30
; RELEVANT RESIDUES: (1)..(154)
PCT-US02-22032-6

Query Match 100.0%; Score 821; DB 1; Length 154;
Best Local Similarity 100.0%; Pred. No. 2.2e-84;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATKAVCVLKGDPVQGIINFEQESNGPVKVWGSIKGLTEGLHGFHVHEFGDNTAGCTS 60
Db 1 MATKAVCVLKGDPVQGIINFEQESNGPVKVWGSIKGLTEGLHGFHVHEFGDNTAGCTS 60
QY 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 120
Db 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 120
QY 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
Db 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154

RESULT 4

US-09-744-016A-2
; Sequence 2, Application US/09744016A
; GENERAL INFORMATION:
; APPLICANT: Dr. Voelkel, Helge
; TITLE OF INVENTION: Method for screening of modulators of calcineurin
; TITLE OF INVENTION: activity
; FILE REFERENCE: A34157PCT
; CURRENT APPLICATION NUMBER: US/09/744,016A
; CURRENT FILING DATE: 2001-10-03
; PRIOR APPLICATION NUMBER: EP98113876
; PRIOR FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-744-016A-2

Query Match 100.0%; Score 821; DB 21; Length 154;
Best Local Similarity 100.0%; Pred. No. 2.2e-84;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATKAVCVLKGDPVQGIINFEQESNGPVKVWGSIKGLTEGLHGFHVHEFGDNTAGCTS 60
Db 1 MATKAVCVLKGDPVQGIINFEQESNGPVKVWGSIKGLTEGLHGFHVHEFGDNTAGCTS 60
QY 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 120
Db 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 120
QY 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
Db 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154

RESULT 5

US-09-744-016A-16
; Sequence 16, Application US/09744016A
; GENERAL INFORMATION:
; APPLICANT: Dr. Voelkel, Helge
; TITLE OF INVENTION: Method for screening of modulators of calcineurin
; TITLE OF INVENTION: activity
; FILE REFERENCE: A34157PCT
; CURRENT APPLICATION NUMBER: US/09/744,016A
; CURRENT FILING DATE: 2001-10-03
; PRIOR APPLICATION NUMBER: EP98113876
; PRIOR FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-744-016A-16

```
Query Match 100.0%; Score 821; DB 21; Length 154;
Best Local Similarity 100.0%; Pred. No. 2.2e-84;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATKAVCVLKGDPVQGIINFEQKESNGPKVWGSIKGLTEGLHGFHVEFGDNTAGCTS 60
   |||||
Db 1 MATKAVCVLKGDPVQGIINFEQKESNGPKVWGSIKGLTEGLHGFHVEFGDNTAGCTS 60

QY 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKDGADVSIEDSVISLSDGHCIIIGRTLTV 120
   |||||
Db 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKDGADVSIEDSVISLSDGHCIIIGRTLTV 120

QY 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
   |||||
Db 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154

RESULT 6
US-09-791-537-38416
; Sequence 38416, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791.537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 38416
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-38416

Query Match 100.0%; Score 821; DB 21; Length 154;
Best Local Similarity 100.0%; Pred. No. 2.2e-84;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATKAVCVLKGDPVQGIINFEQKESNGPKVWGSIKGLTEGLHGFHVEFGDNTAGCTS 60
   |||||
Db 1 MATKAVCVLKGDPVQGIINFEQKESNGPKVWGSIKGLTEGLHGFHVEFGDNTAGCTS 60

QY 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKDGADVSIEDSVISLSDGHCIIIGRTLTV 120
   |||||
Db 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKDGADVSIEDSVISLSDGHCIIIGRTLTV 120

QY 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
   |||||
Db 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154

RESULT 7
US-09-904-987-6
; Sequence 6, Application US/09904987
; GENERAL INFORMATION:
; APPLICANT: Novactyl, Inc.
; TITLE OF INVENTION: Methods and Compositions for Controlling Pathological and Prepath
; TITLE OF INVENTION: Protein Assembly or Aggregation
; FILE REFERENCE: 42108/26146
; CURRENT APPLICATION NUMBER: US/09/904,987
; CURRENT FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 154
; TYPE: PRT
; ORGANISM: homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NCBI ENTREZ / P00441
```

```
; DATABASE ENTRY DATE: 2000-05-30
; RELEVANT RESIDUES: (1)..(154)
US-09-904-987-6

Query Match 100.0%; Score 821; DB 23; Length 154;
Best Local Similarity 100.0%; Pred. No. 2.2e-84;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATKAVCVLKGDPVQGIINFEQKESNGPKVWGSIKGLTEGLHGFHVEFGDNTAGCTS 60
   |||||
Db 1 MATKAVCVLKGDPVQGIINFEQKESNGPKVWGSIKGLTEGLHGFHVEFGDNTAGCTS 60

QY 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKDGADVSIEDSVISLSDGHCIIIGRTLTV 120
   |||||
Db 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKDGADVSIEDSVISLSDGHCIIIGRTLTV 120

QY 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
   |||||
Db 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154

RESULT 8
US-09-760-475-2306
; Sequence 2306, Application US/09760475
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P7249
; CURRENT APPLICATION NUMBER: US/09/760,475
; CURRENT FILING DATE: 2001-01-16
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4122
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2306
; LENGTH: 156
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-760-475-2306

Query Match 100.0%; Score 821; DB 21; Length 156;
Best Local Similarity 100.0%; Pred. No. 2.2e-84;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATKAVCVLKGDPVQGIINFEQKESNGPKVWGSIKGLTEGLHGFHVEFGDNTAGCTS 60
   |||||
Db 3 MATKAVCVLKGDPVQGIINFEQKESNGPKVWGSIKGLTEGLHGFHVEFGDNTAGCTS 62

QY 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKDGADVSIEDSVISLSDGHCIIIGRTLTV 120
   |||||
Db 63 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKDGADVSIEDSVISLSDGHCIIIGRTLTV 122

QY 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
   |||||
Db 123 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 156

RESULT 9
US-08-993-010-2
; Sequence 2, Application US/08993010
; GENERAL INFORMATION:
; APPLICANT: Quan, Stella
; APPLICANT: Valenzuela, Pablo
; APPLICANT: Polito, Alan
; TITLE OF INVENTION: HELICOBACTER PYLORI DIAGNOSTICS
; FILE REFERENCE: 1360.002
; CURRENT APPLICATION NUMBER: US/08/993,010
; CURRENT FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/033,707
; EARLIER FILING DATE: 1996-12-19
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 2
; LENGTH: 668
```

```
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-08-993-010-2

Query Match      100.0%; Score 821; DB 13; Length 668;
Best Local Similarity 100.0%; Pred. No. 1.6e-83;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVHEFGDNTAGCTS 60
   |||||
Db 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVHEFGDNTAGCTS 60

QY 61 AGPHFNPLSRKHGPKDEERHVGDLGNVTADKGDVADVSIEDSVISLSDHCHCIIGRTLTV 120
   |||||
Db 61 AGPHFNPLSRKHGPKDEERHVGDLGNVTADKGDVADVSIEDSVISLSDHCHCIIGRTLTV 120

QY 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
   |||||
Db 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154

RESULT 10
US-09-791-537-91118
; Sequence 91118, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 91118
; LENGTH: 839
; TYPE: PRT
; ORGANISM: synthetic construct
US-09-791-537-91118

Query Match      100.0%; Score 821; DB 21; Length 839;
Best Local Similarity 100.0%; Pred. No. 2.3e-83;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVHEFGDNTAGCTS 60
   |||||
Db 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVHEFGDNTAGCTS 60

QY 61 AGPHFNPLSRKHGPKDEERHVGDLGNVTADKGDVADVSIEDSVISLSDHCHCIIGRTLTV 120
   |||||
Db 61 AGPHFNPLSRKHGPKDEERHVGDLGNVTADKGDVADVSIEDSVISLSDHCHCIIGRTLTV 120

QY 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
   |||||
Db 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154

RESULT 11
US-09-881-654-4
; Sequence 4, Application US/09881654
; GENERAL INFORMATION:
; APPLICANT: CHIEN, David Y.
; APPLICANT: ARCANGEL, Philip
; APPLICANT: TANDESKE, Laura
; APPLICANT: GEORGE-NASCIEMENTO, Carlos
; APPLICANT: COLT, Doris
; APPLICANT: MEDINA-SELBY, Angelica
; TITLE OF INVENTION: IMMUNOASSAYS FOR ANTI-HCV ANTIBODIES
; FILE REFERENCE: 2302-17039 / PPI7039.002
; CURRENT APPLICATION NUMBER: US/09/881,654
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: 60/212,082
```

```
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/280,811
; PRIOR FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: 60/280,867
; PRIOR FILING DATE: 2001-04-02
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1099
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: MEFA 7.1
US-09-881-654-4

Query Match      100.0%; Score 821; DB 22; Length 1099;
Best Local Similarity 100.0%; Pred. No. 3.3e-83;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVHEFGDNTAGCTS 60
   |||||
Db 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVHEFGDNTAGCTS 60

QY 61 AGPHFNPLSRKHGPKDEERHVGDLGNVTADKGDVADVSIEDSVISLSDHCHCIIGRTLTV 120
   |||||
Db 61 AGPHFNPLSRKHGPKDEERHVGDLGNVTADKGDVADVSIEDSVISLSDHCHCIIGRTLTV 120

QY 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
   |||||
Db 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154

RESULT 12
US-07-768-912-13
; Sequence 13, Application US/07768912
; GENERAL INFORMATION:
; APPLICANT: Smith, Geoffrey
; TITLE OF INVENTION: VACCINIA VECTORS, VACCINIA GENES AND
; TITLE OF INVENTION: EXPRESSION PRODUCTS THEREOF
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Walter H. Dreger
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/07/768,912
; APPLICATION NUMBER: US/07/768,912
; FILING DATE: 19911112
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Walter H.
; REGISTRATION NUMBER: 24,190
; REFERENCE/DOCKET NUMBER: A-55557
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 153 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
us-07-768-912-13

Query Match      99.4%; Score 816; DB 3; Length 153;
```

Best Local Similarity 100.0%; Pred. No. 7.9e-84;
Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVEFGDNTAGCTSA 61
DB 1 ATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVEFGDNTAGCTSA 60
QY 62 GPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSDSVISLSDGHCIIIGRTLTVVH 121
DB 61 GPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSDSVISLSDGHCIIIGRTLTVVH 120
QY 122 EKADDLGKGGNEESTKTGNAGSLRACGVIGIAQ 154
DB 121 EKADDLGKGGNEESTKTGNAGSLRACGVIGIAQ 153

RESULT 13

US-09-744-016A-14

; Sequence 14, Application US/09744016A

; GENERAL INFORMATION:

; APPLICANT: Dr. Voelkel, Helge

; TITLE OF INVENTION: Method for screening of modulators of calcineurin

; FILE REFERENCE: A34157PCT

; CURRENT APPLICATION NUMBER: US/09/744,016A

; PRIOR FILING DATE: 2001-10-03

; PRIOR APPLICATION NUMBER: EP98113876

; PRIOR FILING DATE: 1998-07-22

; NUMBER OF SEQ ID NOS: 35

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 14

; LENGTH: 153

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-744-016A-14

Query Match 99.4%; Score 816; DB 21; Length 153;

Best Local Similarity 100.0%; Pred. No. 7.9e-84;

Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVEFGDNTAGCTSA 61
DB 1 ATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVEFGDNTAGCTSA 60
QY 62 GPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSDSVISLSDGHCIIIGRTLTVVH 121
DB 61 GPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSDSVISLSDGHCIIIGRTLTVVH 120
QY 122 EKADDLGKGGNEESTKTGNAGSLRACGVIGIAQ 154
DB 121 EKADDLGKGGNEESTKTGNAGSLRACGVIGIAQ 153

RESULT 14

US-10-018-770-1

; Sequence 1, Application US/10018770

; GENERAL INFORMATION:

; APPLICANT: LIT Institute Co., Ltd., and Seikagaku Corporation

; TITLE OF INVENTION: Pharmaceutical composition containing lecithinized-superoxide dis

; FILE REFERENCE:

; CURRENT APPLICATION NUMBER: US/10/018,770

; CURRENT FILING DATE: 2001-12-17

; NUMBER OF SEQ ID NOS: 1

; SEQ ID NO 1

; LENGTH: 153

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-018-770-1

Query Match

Best Local Similarity 99.4%; Score 816; DB 24; Length 153;

Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVEFGDNTAGCTSA 61
DB 1 ATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVEFGDNTAGCTSA 60
QY 62 GPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSDSVISLSDGHCIIIGRTLTVVH 121
DB 61 GPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSDSVISLSDGHCIIIGRTLTVVH 120
QY 122 EKADDLGKGGNEESTKTGNAGSLRACGVIGIAQ 154
DB 121 EKADDLGKGGNEESTKTGNAGSLRACGVIGIAQ 153

RESULT 15

US-09-791-537-151133

; Sequence 151133, Application US/09791537

; GENERAL INFORMATION:

; APPLICANT: Bionomix, Inc.

; APPLICANT: Debe, Derek

; APPLICANT: Danzer, Joseph

; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME

; FILE REFERENCE: 261/210

; CURRENT APPLICATION NUMBER: US/09/791,537

; CURRENT FILING DATE: 2001-02-22

; NUMBER OF SEQ ID NOS: 153055

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 151133

; LENGTH: 154

; TYPE: PRT

; ORGANISM: pdb 1SPDA

; NAME/KEY: misc_feature

; LOCATION: (1)..(1)

; OTHER INFORMATION: X is an unknown amino acid

US-09-791-537-151133

Query Match 99.4%; Score 816; DB 21; Length 154;

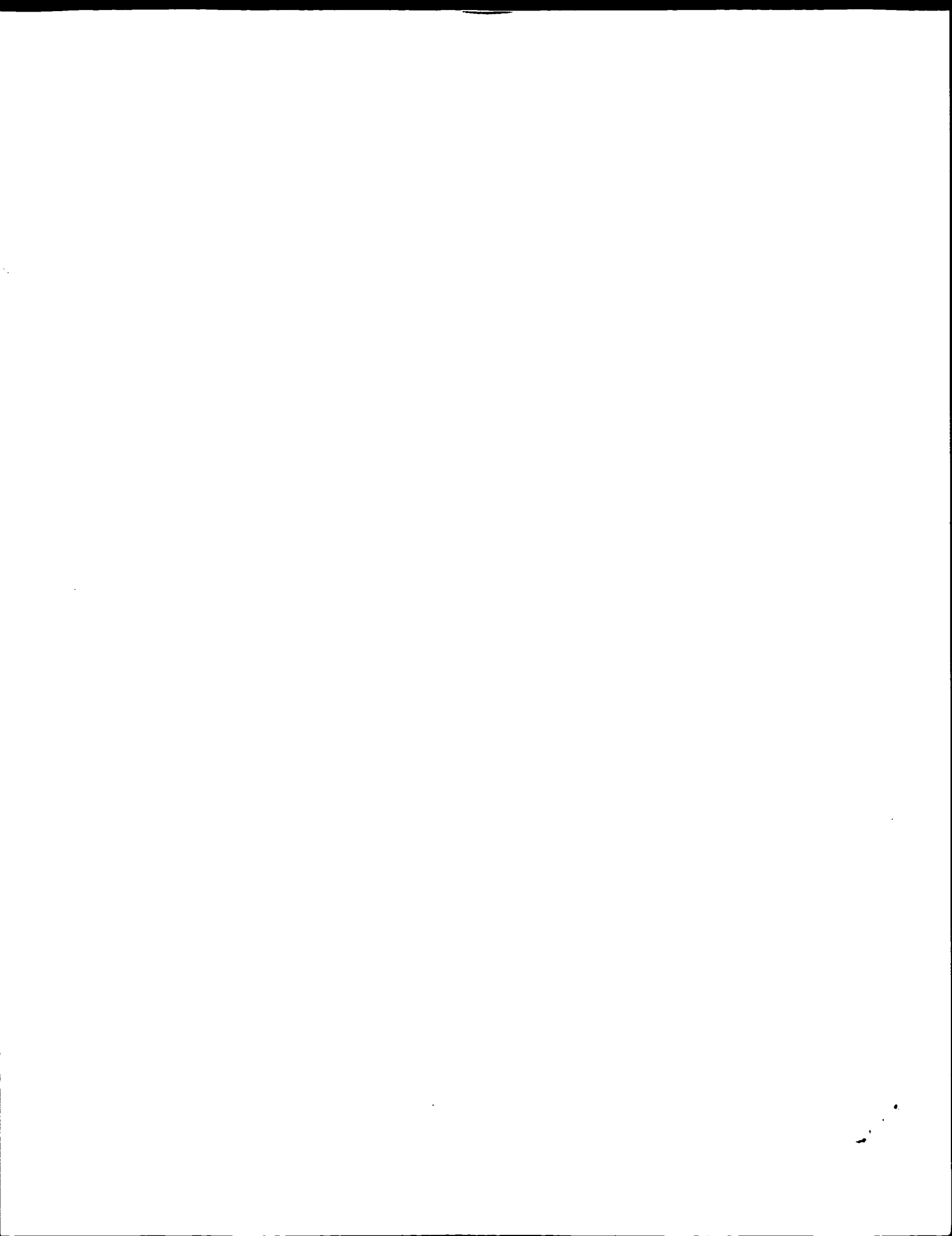
Best Local Similarity 100.0%; Pred. No. 8e-84;

Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVEFGDNTAGCTSA 61
DB 2 ATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVEFGDNTAGCTSA 61
QY 62 GPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSDSVISLSDGHCIIIGRTLTVVH 121
DB 62 GPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSDSVISLSDGHCIIIGRTLTVVH 121
QY 122 EKADDLGKGGNEESTKTGNAGSLRACGVIGIAQ 154
DB 122 EKADDLGKGGNEESTKTGNAGSLRACGVIGIAQ 154

Search completed: January 28, 2003, 14:37:27

Job time : 41.9197 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:16:13 ; Search time 4.50096 Seconds
(without alignments)
2701.343 Million cell updates/sec

Title: US-09-904-987-6
Perfect score: 821
Sequence: 1 MATKAVCVLKGDGPVQGIIN.....STKGNAGSLRACGVIGIAQ 154

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 394613
Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pap:*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pap:*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pap:*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pap:*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pap:*
6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pap:*
7: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
1	821	100.0	154	1	PCT-US02-35563-20	Sequence 20, Appl
2	821	100.0	236	1	PCT-US02-33121-46	Sequence 46, Appl
3	821	100.0	236	6	US-10-272-459-46	Sequence 46, Appl
4	821	100.0	352	1	PCT-US02-33121-45	Sequence 45, Appl
5	821	100.0	352	6	US-10-272-459-45	Sequence 45, Appl
6	821	100.0	382	1	PCT-US02-33121-47	Sequence 47, Appl
7	821	100.0	382	6	US-10-272-459-47	Sequence 47, Appl
8	821	100.0	652	1	PCT-US02-33121-48	Sequence 48, Appl
9	821	100.0	652	6	US-10-272-459-48	Sequence 48, Appl
10	682	83.1	127	5	US-09-513-999C-5289	Sequence 5289, Ap
11	644	78.4	128	5	US-09-724-676-81086	Sequence 81086, A
12	644	78.4	128	5	US-09-724-676-81097	Sequence 81097, A
13	644	78.4	128	5	US-09-724-676-81107	Sequence 81107, A
14	644	78.4	128	5	US-09-724-676-81112	Sequence 81112, A
15	644	78.4	128	5	US-09-724-676A-81086	Sequence 81086, A
16	644	78.4	128	5	US-09-724-676A-81107	Sequence 81107, A
17	528.5	64.4	150	5	US-09-724-676A-81107	Sequence 81107, A
18	528.5	64.4	150	5	US-09-724-676A-81112	Sequence 81112, A
19	449.5	54.8	88	5	US-09-724-676-81111	Sequence 81111, A
20	449.5	54.8	88	5	US-09-724-676A-81111	Sequence 81111, A
21	447.5	54.5	88	5	US-09-724-676-81089	Sequence 81089, A
22	447.5	54.5	88	5	US-09-724-676-81100	Sequence 81100, A
23	447.5	54.5	88	5	US-09-724-676A-81089	Sequence 81089, A
24	447.5	54.5	88	5	US-09-724-676A-81100	Sequence 81100, A
25	442	53.8	80	5	US-09-724-676-81087	Sequence 81087, A
26	442	53.8	80	5	US-09-724-676-81088	Sequence 81088, A

27	442	53.8	80	5	US-09-724-676-81098	Sequence 81098, A
28	442	53.8	80	5	US-09-724-676-81099	Sequence 81099, A
29	442	53.8	80	5	US-09-724-676-81108	Sequence 81108, A
30	442	53.8	80	5	US-09-724-676-81109	Sequence 81109, A
31	442	53.8	80	5	US-09-724-676A-81087	Sequence 81087, A
32	442	53.8	80	5	US-09-724-676A-81088	Sequence 81088, A
33	442	53.8	80	5	US-09-724-676A-81098	Sequence 81098, A
34	442	53.8	80	5	US-09-724-676A-81099	Sequence 81099, A
35	442	53.8	80	5	US-09-724-676A-81108	Sequence 81108, A
36	442	53.8	80	5	US-09-724-676A-81109	Sequence 81109, A
37	365	44.5	218	5	US-09-724-676-78095	Sequence 78095, A
38	365	44.5	218	5	US-09-724-676A-78095	Sequence 78095, A
39	305.5	37.2	80	5	US-09-724-676-81101	Sequence 81101, A
40	305.5	37.2	80	5	US-09-724-676A-81101	Sequence 81101, A
41	303	36.9	66	5	US-09-724-676-81096	Sequence 81096, A
42	303	36.9	66	5	US-09-724-676A-81096	Sequence 81096, A
43	303	36.9	69	5	US-09-724-676-81090	Sequence 81090, A
44	303	36.9	69	5	US-09-724-676-81110	Sequence 81110, A
45	303	36.9	69	5	US-09-724-676A-81090	Sequence 81090, A

ALIGNMENTS

RESULT 1

PCT-US02-35563-20
; Sequence 20, Application PC/TUS0235563
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies, Inc
; TITLE OF INVENTION: REGULATED PROSTATE CANCER GENES
; FILE REFERENCE: Oct 90 803 PCT
; CURRENT APPLICATION NUMBER: PCT/US02/35563
; CURRENT FILING DATE: 2002-11-07
; PRIOR APPLICATION NUMBER: 60/331,042
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: 60/331,041
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: 60/340,251
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 60/344,791
; PRIOR FILING DATE: 2002-01-07
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-35563-20

Query Match 100.0%; Score 821; DB 1; Length 154;
Best Local Similarity 100.0%; Pred. No. 3.8e-61;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MATKAVCVLKGDGPVQGIINFEOKESNGPVKVMGSIKGLTEGLHGFHVHEFGDNTAGCT	60
Db	1	MATKAVCVLKGDGPVQGIINFEOKESNGPVKVMGSIKGLTEGLHGFHVHEFGDNTAGCT	60
QY	61	AGPHFNPLSRKKGKDEERHVGDLGNVTADKDGADVSDSVISLSDGHCIIIGRTLVV	120
Db	61	AGPHFNPLSRKKGKDEERHVGDLGNVTADKDGADVSDSVISLSDGHCIIIGRTLVV	120
QY	121	HEKADDLGKGNNESTKTGNAGSLRACGVIGIAQ	154
Db	121	HEKADDLGKGNNESTKTGNAGSLRACGVIGIAQ	154

RESULT 2

PCT-US02-33121-46
; Sequence 46, Application PC/TUS0233121
; GENERAL INFORMATION:
; APPLICANT: CHIRON CORPORATION et al.
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
; TITLE OF INVENTION: PROTEINS AND USES THEREOF

```
; FILE REFERENCE: PP17955.003 / 2301-17955.40
; CURRENT APPLICATION NUMBER: PCT/US02/33121
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 46
; TYPE: PRT
; LENGTH: 236
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
; OTHER INFORMATION: of 24.9 kDa
PCT-US02-33121-46

Query Match          100.0%; Score 821; DB 1; Length 236;
Best Local Similarity 100.0%; Pred. No. 6.3e-61;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVHFEFGDNTAGCTS 60
   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVHFEFGDNTAGCTS 60

QY 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDHCHIIIGRTLTV 120
   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDHCHIIIGRTLTV 120

QY 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154

RESULT 3
US-10-272-459-46
; Sequence 46, Application US/10272459
; GENERAL INFORMATION:
; APPLICANT: PICHUANES, Sergio
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
; FILE REFERENCE: PP17955.002 / 2301-17955
; CURRENT APPLICATION NUMBER: US/10/272,459
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 46
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
; OTHER INFORMATION: of 24.9 kDa
US-10-272-459-46

Query Match          100.0%; Score 821; DB 6; Length 236;
Best Local Similarity 100.0%; Pred. No. 6.3e-61;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVHFEFGDNTAGCTS 60
   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVHFEFGDNTAGCTS 60

QY 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSTEDSVISLSDHCHIIIGRTLTV 120
   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSTEDSVISLSDHCHIIIGRTLTV 120

QY 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154

RESULT 4
PCT-US02-33121-45
; Sequence 45, Application PC/TUS0233121
; GENERAL INFORMATION:
; APPLICANT: CHIRON CORPORATION et al.
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
; FILE REFERENCE: PP17955.003 / 2301-17955.40
; CURRENT APPLICATION NUMBER: PCT/US02/33121
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 45
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
; OTHER INFORMATION: of 38.8 kDa
PCT-US02-33121-45

Query Match          100.0%; Score 821; DB 1; Length 352;
Best Local Similarity 100.0%; Pred. No. 9.9e-61;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVHFEFGDNTAGCTS 60
   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVHFEFGDNTAGCTS 60

QY 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDHCHIIIGRTLTV 120
   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDHCHIIIGRTLTV 120

QY 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154

RESULT 5
US-10-272-459-45
; Sequence 45, Application US/10272459
; GENERAL INFORMATION:
; APPLICANT: PICHUANES, Sergio
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
; FILE REFERENCE: PP17955.002 / 2301-17955
; CURRENT APPLICATION NUMBER: US/10/272,459
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 45
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
; OTHER INFORMATION: of 38.8 kDa
US-10-272-459-45

Query Match          100.0%; Score 821; DB 6; Length 352;
Best Local Similarity 100.0%; Pred. No. 9.9e-61;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVHFEFGDNTAGCTS 60
   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVHFEFGDNTAGCTS 60

QY 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDHCHIIIGRTLTV 120
   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDHCHIIIGRTLTV 120

QY 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154

RESULT 6
```

```

PCT-US02-33121-47
; Sequence 47, Application PC/TUS0233121
; GENERAL INFORMATION:
; APPLICANT: CHIRON CORPORATION et al.
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
; TITLE OF INVENTION: PROTEINS AND USES THEREOF
; FILE REFERENCE: PP17955.003 / 2301-17955.40
; CURRENT APPLICATION NUMBER: PCT/US02/33121
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 47
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
; OTHER INFORMATION: of 41 kDa
PCT-US02-33121-47

```

```

RESULT 7
US-10-272-459-47
; Sequence 47, Application US/10272459
; GENERAL INFORMATION:
; APPLICANT: PICHUANTES, Sergio
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
; FILE REFERENCE: PP17955.002 / 2301-17955
; CURRENT APPLICATION NUMBER: US/10/272,459
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 47
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
; OTHER INFORMATION: of 41 kDa
US-10-272-459-47

```

```

RESULT 8
PCT-US02-33121-48
; Sequence 48, Application PC/TUS0233121
; GENERAL INFORMATION:
; APPLICANT: CHIRON CORPORATION et al.
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
; TITLE OF INVENTION: PROTEINS AND USES THEREOF
; FILE REFERENCE: PP17955.003 / 2301-17955.40
; CURRENT APPLICATION NUMBER: PCT/US02/33121
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 48
; LENGTH: 652
; TYPE: PRY
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
; OTHER INFORMATION: of human superoxide dismutase fused with the HAV
; OTHER INFORMATION: nonstructural protein
PCT-US02-33121-48

```

```

RESULT 9
US-10-272-459-48
; Sequence 48, Application US/10272459
; GENERAL INFORMATION:
; APPLICANT: PICHUANES, Sergio
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
; TITLE OF INVENTION: PROTEINS AND USES THEREOF
; FILE REFERENCE: PPI7955.002 / 2301-17955
; CURRENT APPLICATION NUMBER: US/10/272.459
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 48
; LENGTH: 652
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
; OTHER INFORMATION: of human superoxide dismutase fused with the HAV
; OTHER INFORMATION: nonstructural protein
US-10-272-459-48

```

	Query Match	100.0%	Score 821;	DB 6;	Length 652;
	Best Local Similarity	100.0%	Pred. No. 2e-60;		
	Matches 154;	Conservative	0;	Mismatches 0;	Indels 0; Caps 0
Qy	1	MATKAVCVLKGDPVGGIINFQKESNGPVKVMGSIKGLTEGLHGFVHFEFGONTAGCTS	60		
Db	1	MATKAVCVLKGDPVGGIINFQKESNGPVKVMGSIKGLTEGLHGFVHFEFGONTAGCTS	60		
Qy	61	AGPHFNPLSRKHGGPDEERHVRGDLGNVTADKGDVADYSIEDSVISLSGHCIIIGRTLTV	120		

D6	61	AGPHFPLSRKHKGPKDEERHVGDLGNVTADKGVADVSI	SVLSGDHCIGRTL	V120
Q7	121	HEKADDLGKGNEESTTKTNAGSRLACGVIGIAQ		154
D6	121	HEKADDLGKGNEESTTKTNAGSRLACGVIGIAQ		154

```

RESULT 10
US-09-513-999C-5289
; Sequence 5289, Application US/09513999C
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59 US2 REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 5289
; LENGTH: 127
; TYPE: prt
; ORGANISM: Homo sapiens
US-09-513-999C-5289

```

```

RESULT 11
US-09-724-676-81086
; Sequence 81086, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 81086
; LENGTH: 128
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-81086

```

```

Query Match      78.4%; Score 644; DB 5; Length 128;
Best Local Similarity 100.0%; Pred. No. 1.6e-46;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MATKAVCVLKDGPPVQGIINFEQKESNPVKVWGSIKGLTEGLHGFHVHFGDNTAGCTS 60
        |||||
Db       1 MATKAVCVLKDGPPVQGIINFEQKESNPVKVWGSIKGLTEGLHGFHVHFGDNTAGCTS 60
        |||||

QY      61 ACPHFNPLSRKHGGPKDEERHVGDLGNVTKDQGVADVSIEDSVISLSDGHCIIIGRTLW 120
        |||||
Db       61 ACPHFNPLSRKHGGPKDEERHVGDLGNVTKDQGVADVSIEDSVISLSDGHCIIIGRTLW 120
        |||||

```

```

RESULT 12
US-09-724-674-81097
; Sequence 81097, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: CompuGen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 CompuGen
; CURRENT APPLICATION NUMBER: US/09/724.676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 81097
; LENGTH: 128
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-81097

```

```

RESULT 13
US-09-724-676-81107
; Sequence 81107, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: CompuGen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 CompuGen
; CURRENT APPLICATION NUMBER: US/09/724, 676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 81107
; LENGTH: 128
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-81107

```

```

RESULT 14
; US-09-724-676A-81086
; Sequence 81086, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: CompuGen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 CompuGen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2

```

; SEQ ID NO 81086
; LENGTH: 128
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-81086

Query Match 78.4%; Score 644; DB 5; Length 128;
Best Local Similarity 100.0%; Pred. No. 1.6e-46;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVHFEFGDNTAGCTS 60
Db 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVHFEFGDNTAGCTS 60
QY 61 AGPHFNPLSRKHGKGPDEERHVDLGNVTADKDGVDVSDVSIKGLTEGLHGFHVHFEFGDNTAGCTS 120
Db 61 AGPHFNPLSRKHGKGPDEERHVDLGNVTADKDGVDVSDVSIKGLTEGLHGFHVHFEFGDNTAGCTS 120

RESULT 15

US-09-724-676A-81097
; Sequence 81097, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 CompuGen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 81097
; LENGTH: 128
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-81097

Query Match 78.4%; Score 644; DB 5; Length 128;
Best Local Similarity 100.0%; Pred. No. 1.6e-46;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVHFEFGDNTAGCTS 60
Db 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVHFEFGDNTAGCTS 60
QY 61 AGPHFNPLSRKHGKGPDEERHVDLGNVTADKDGVDVSDVSIKGLTEGLHGFHVHFEFGDNTAGCTS 120
Db 61 AGPHFNPLSRKHGKGPDEERHVDLGNVTADKDGVDVSDVSIKGLTEGLHGFHVHFEFGDNTAGCTS 120

Search completed: January 28, 2003, 14:39:22
Job time : 6.50096 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:10:57 : Search time 3.19694 Seconds
(without alignments)
1417.333 Million cell updates/sec

Title: US-09-904-987-6

Perfect score: 821

Sequence: 1 MATKAVCVLKGDGPVQGIIN.....STKGNAGSLACGVIGIAQ 154

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*

2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*

3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*

4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*

5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*

6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	821	100.0	154	4	US-09-126-109-4
2	821	100.0	618	1	US-08-668-381A-5
3	821	100.0	1021	1	US-07-910-760-12
4	821	100.0	1021	1	US-08-440-519-12
5	821	100.0	1021	4	US-08-440-549-12
6	816	99.4	153	4	US-08-679-493A-204
7	816	99.4	153	6	5290690-6
8	812	98.9	154	6	5290690-5
9	807	98.3	152	2	US-08-722-050-7
10	801.5	97.6	152	6	5196335-1
11	797	97.1	841	1	US-08-350-884-86
12	797	97.1	841	1	US-08-709-173-86
13	797	97.1	841	2	US-08-709-177-86
14	772	94.0	144	6	5252476-5
15	758	92.3	152	6	5171680-3
16	744	90.6	154	2	US-08-023-980B-43
17	744	90.6	154	2	US-08-486-953A-51
18	684.5	83.4	152	4	US-08-679-493A-203
19	673	82.0	153	4	US-08-679-493A-202
20	670	81.6	153	4	US-08-679-493A-201
21	662	80.6	153	4	US-08-679-493A-207
22	655.5	79.8	151	2	US-08-722-050-10
23	650.5	79.2	152	4	US-08-679-493A-206
24	647	78.8	151	4	US-09-202-832-16
25	647	78.8	151	4	US-08-679-493A-205
26	638	77.7	150	2	US-08-722-050-4
27	638	77.7	151	6	5290690-7

ALIGNMENTS

RESULT 1

US-09-126-109-4

; Sequence 4, Application US/09126109

; Patent No. 6171856

; GENERAL INFORMATION:

; APPLICANT: Thigpen, Anice

; APPLICANT: Holmeier, Hans-Ewald

; APPLICANT: Newgard, Christopher B.

; APPLICANT: Unger, Roger H.

; APPLICANT: Shimabukuro, Michio

; APPLICANT: Chen, Guaxun

; APPLICANT: Rhodes, Christopher J.

; APPLICANT: Hugl, Sigrun R.

; APPLICANT: Cousin, Sharon

; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING

; TITLE OF INVENTION: TO NO-MEDIATED CYTOTOXICITY

; NUMBER OF SEQUENCES: 20

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Arnold, White & Durkee

; CITY: Houston

; STATE: Texas

; COUNTRY: USA

; ZIP: 77210

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09126,109

; FILING DATE: 30-JUL-1998

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 60/055,092

; FILING DATE: 30-JUL-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US Unknown

; FILING DATE: 03-MAR-1998

; ATTORNEY/AGENT INFORMATION:

; NAME: McMillian, Nabeela R.

; REGISTRATION NUMBER: P-43,363

; REFERENCE/DOCKET NUMBER: UTSD:560

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (512) 418-3000

; TELEFAX: (512) 474-7577

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 154 amino acids

; TYPE: amino acid

Sequence 209, App
Sequence 208, App
Sequence 200, App
Sequence 11, Appl
Sequence 199, App
Sequence 3, Appl
Sequence 6, Appl
Sequence 9, Appl
Sequence 192, App
Sequence 210, App
Sequence 191, App
Sequence 190, App
Sequence 197, App
Sequence 195, App
Sequence 196, App
Sequence 198, App
Sequence 5, Appl
Sequence 8, Appl

589.5 71.8 166 4 US-08-679-493A-209
558.5 68.0 152 4 US-08-679-493A-208
527.5 64.3 149 4 US-08-679-493A-200
522 63.6 150 2 US-08-722-050-11
521 63.5 150 4 US-08-679-493A-199
492 59.9 153 2 US-08-722-050-3
483 58.8 151 2 US-08-722-050-9
475.5 57.9 151 4 US-08-679-493A-192
471.5 57.4 151 4 US-08-679-493A-210
469.5 57.2 151 4 US-08-679-493A-191
464.5 56.6 201 4 US-08-679-493A-190
453.5 56.5 202 4 US-08-679-493A-197
455.5 55.5 218 4 US-08-679-493A-195
453.5 55.2 217 4 US-08-679-493A-196
448 54.6 221 4 US-08-679-493A-198
444.5 54.1 151 2 US-08-722-050-5
444.5 54.1 151 2 US-08-722-050-8

STRANDEDNESS:
TOPOLOGY: linear
US-09-126-109-4

Query Match 100.0%; Score 821; DB 4; Length 154;
Best Local Similarity 100.0%; Pred. No. 7.7e-85;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVEFGDNTAGCTS 60
DB 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVEFGDNTAGCTS 60
QY 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKDGADVADSVISSLSDHCHIIIGRTLTV 120
DB 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKDGADVADSVISSLSDHCHIIIGRTLTV 120
QY 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
DB 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154

RESULT 2

US-08-668-381A-5
Sequence 5, Application US/08668381A
Patent No. 5780024
GENERAL INFORMATION:
APPLICANT: Brown, Robert H.
APPLICANT: Fishman, Paul S.
APPLICANT: Francis, Jonathan W.
APPLICANT: Hosler, Betsy A.
TITLE OF INVENTION: SUPEROXIDE DISMUTASE/TETANUS TOXIN
TITLE OF INVENTION: FRAGMENT C HYBRID PROTEIN
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/668,381A
FILING DATE: 21-JUN-1996
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/000,473
FILING DATE: 23-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,164
REFERENCE/DOCKET NUMBER: 00786/269001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 618 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-668-381A-5

Query Match 100.0%; Score 821; DB 1; Length 618;
Best Local Similarity 100.0%; Pred. No. 5e-84;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVEFGDNTAGCTS 60

DB 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVEFGDNTAGCTS 60
QY 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKDGADVADSVISSLSDHCHIIIGRTLTV 120
DB 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKDGADVADSVISSLSDHCHIIIGRTLTV 120
QY 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
DB 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
RESULT 3
US-07-910-760-12
Sequence 12, Application US/07910760
Patent No. 5683864
GENERAL INFORMATION:
APPLICANT: Houghton, Michael
APPLICANT: Choo, Qui-Lim
APPLICANT: Kuo, George
TITLE OF INVENTION: Combinations of Hepatitis C virus (HCV)
TITLE OF INVENTION: Antigens for Use in Immunoassays for Anti-HCV Antibodies
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: P.O. Box 8097 (Int. Prop. R-440)
CITY: Emeryville
STATE: CA
COUNTRY: U.S.A.
ZIP: 94662-8097
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/910,760
FILING DATE: 07-JUL-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Blackburn Esq., Robert P.
REGISTRATION NUMBER: 30,447
REFERENCE/DOCKET NUMBER: 0101,002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 601-2702
TELEFAX: (510) 655-3542
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1021 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-910-760-12

Query Match 100.0%; Score 821; DB 1; Length 1021;
Best Local Similarity 100.0%; Pred. No. 9.9e-84;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVEFGDNTAGCTS 60
DB 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVEFGDNTAGCTS 60
QY 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKDGADVADSVISSLSDHCHIIIGRTLTV 120
DB 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKDGADVADSVISSLSDHCHIIIGRTLTV 120
QY 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
DB 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154

RESULT 4
US-08-440-519-12

Sequence 12, Application US/08440519
Patent No. 5712087
GENERAL INFORMATION:
APPLICANT: Houghton, Michael
APPLICANT: Choo, Qui-Lim
APPLICANT: Kuo, George
TITLE OF INVENTION: Combinations of Hepatitis C virus (HCV)
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: P.O. Box 8097 (Int. Prop. R-440)
CITY: Emeryville
STATE: CA
COUNTRY: U.S.A.
ZIP: 94662-8097
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/440,519
FILING DATE: 12-MAY-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/910,760
FILING DATE: 07-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Blackburn Esq., Robert P.
REGISTRATION NUMBER: 30,447
REFERENCE/DOCKET NUMBER: 0101.002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 601-2702
TELEFAX: (510) 655-3542
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1021 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-440-519-12

Query Match 100.0%; Score 821; DB 1; Length 1021;
Best Local Similarity 100.0%; Pred. No. 9.9e-84;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVHFGDNTAGCTS 60
|||||
DB 1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVHFGDNTAGCTS 60
|||||

QY 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLIV 120
|||||
DB 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLIV 120
|||||

QY 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
|||||
DB 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
|||||

RESULT 5
US-08-440-549-12
Sequence 12, Application US/08440549
Patent No. 6312889
GENERAL INFORMATION:
APPLICANT: Houghton, Michael
APPLICANT: Choo, Qui-Lim
APPLICANT: Kuo, George
TITLE OF INVENTION: Combinations of Hepatitis C virus (HCV)
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation

STREET: P.O. Box 8097 (Int. Prop. R-440)
CITY: Emeryville
STATE: CA
COUNTRY: U.S.A.
ZIP: 94662-8097
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/440,549
FILING DATE: 12-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/910,760
FILING DATE: 07-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Blackburn Esq., Robert P.
REGISTRATION NUMBER: 30,447
REFERENCE/DOCKET NUMBER: 0101.002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 601-2702
TELEFAX: (510) 655-3542
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1021 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-440-549-12

Query Match 100.0%; Score 821; DB 4; Length 1021;
Best Local Similarity 100.0%; Pred. No. 9.9e-84;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVHFGDNTAGCTS 60
|||||
DB 1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVHFGDNTAGCTS 60
|||||

QY 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLIV 120
|||||
DB 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLIV 120
|||||

QY 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
|||||
DB 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
|||||

RESULT 6
US-08-679-493A-204
Sequence 204, Application US/08679493A
Patent No. 6303295
GENERAL INFORMATION:
APPLICANT: Taylor, Ethan W.
TITLE OF INVENTION: SELENOPROTEINS, CODING SEQUENCES AND METHODS
FILE REFERENCE: 55-95
CURRENT APPLICATION NUMBER: US/08/679,493A
CURRENT FILING DATE: 1996-07-12
PRIOR APPLICATION NUMBER: 60/001203
PRIOR FILING DATE: 1995-07-14
PRIOR APPLICATION NUMBER: 60/003,112
PRIOR FILING DATE: 1995-09-01
NUMBER OF SEQ ID NOS: 216
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 204
LENGTH: 153
TYPE: PRT
ORGANISM: Homo sapiens
US-08-679-493A-204

Query Match 99.4%; Score 816; DB 4; Length 153;
Best Local Similarity 100.0%; Pred. No. 2.8e-84;

Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

[illegible]

RESULT 7

```

5290690-6
; Patent No. 5290690
; APPLICANT: MRABET, NADIR; LASTERS, IG
; MATTHYSSENS, GASTON; WODAK, SHOSHANA; OUAH
; TITLE OF INVENTION: METHODS AND MEAN
; STABILITY OF PROTEINS
; NUMBER OF SEQUENCES: 22
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/398,706
; FILING DATE: 25-AUG-1989
; SEQ ID NO: 6:
; LENGTH: 153
5290690-6

```

```
Query Match          99.4%; Score 816; DB 6; Length 153;
Best Local Similarity 100.0%; Pred. No. 2.8e-84;
Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY	2	ATKAVCVLKGDPVQGIINFEQESKSNPKVWGSIKGLTEGLGHGFHVFHFGDNTAGCTSA	61
DB	1	ATKAVCVLKGDPVQGIINFEQESKSNPKVWGSIKGLTEGLGHGFHVFHFGDNTAGCTSA	60
QY	62	GFHFNPLSRKGGPKDERHVGDLGNWTADKQGVADVISEDVSLSDGHCIIIGRTLWVH	121
DB	61	GFHFNPLSRKGGPKDERHVGDLGNWTADKQGVADVISEDVSLSDGHCIIIGRTLWVH	120
QY	122	EKADDLGKGGNEESTKTGNAGSLACGVIGIAQ	154
DB	121	EKADDLGKGGNEESTKTGNAGSLACGVIGIAQ	153

RESULT 8

5290690-5 ; Patent No. 5290690
; APPLICANT: MRABET, NADIR; LASTERS, IG
; MATTHYSSENS, GASTON; WODAK, SHOSHANA; QUAX
; TITLE OF INVENTION: METHODS AND MEANS
; STABILITY OF PROTEINS
; NUMBER OF SEQUENCES: 22
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/398,706
; FILING DATE: 25-AUG-1989
; SEQ ID NO:5
; LENGTH: 154
5290690-5

```
Query Match          98.9%; Score 812; DB 6; Length 154;
Best Local Similarity 99.4%; Pred. No. 7.9e-84;
Matches 153: Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

[illegible]

QY 121 HEKADDLGKGNEESTKTGNAGSRLACGVIGIAQ 154
|||
Db 121 HEKADDLGKGNEESTKTGNAGSRLACGVIGIAQ 154

RESULT 9

```

US-08-722-050-7
; Sequence 7, Application US/08722050
; Patent No. 5871729
; GENERAL INFORMATION:
; APPLICANT: YU, GUO-LIANG
; APPLICANT: ROSEN, CRAIG A.
; APPLICANT: FRASER, CLAIRE M.
; APPLICANT: GOCAYNE, JEANNINE D.
; TITLE OF INVENTION: SUPEROXIDE DISMUTASE-4
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, N.W., SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/722,050
; FILING DATE: 23-JAN-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/235,757
; FILING DATE: 11-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: STEFFE, ERIC K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.1020001/EKS/AJK
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 152 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-722-050-7

```

Query Match	98.3%	Score 807;	DB 2;	Length 152;
Best Local Similarity	99.3%	Pred. No. 2.8e-83;		
Matches 151;	Conservative	0;	Mismatches 1;	Gaps 0;
			Indels	0;

Qy	3	TRAVCVLKGDPVQGIINFEQKESNGPVKVWGSIKGLTEGLHGFHVHFEFGDNTAGCTSAG	62
Db	1	TRAVCVLKGDPVQGIINFEQKESNGPVKVWGSIKGLTEGLHGFHVHFEFGDNTAGCTSAG	60
Qy	63	PHFNPLSRKHGPKDDEERHVGDLGNVTADKGVADVSIEDSVISLSGDHCIIIGRTLVVHE	122
Db	61	PHFNPLSRKHGPKDDEERHVGDLGNVTADKGVADVSIEDSVISLSGDHCIIIGRTLVVHE	120
Qy	123	KADDLKGKGNNEESTKTGNAGSRACGVIGTQ	154
Db	121	KADDLKGKGNNEESTKTGNAGSRACGVIGTQ	152

RESULT 10

5196335-1
; Patent No. 5196335
; APPLICANT: GRONER, YORAM

;; TITLE OF INVENTION: HUMAN SUPEROXIDE DISMUTASE CDNA

;; NUMBER OF SEQUENCES: 2

;; CURRENT APPLICATION DATA:

;; FILING DATE: 2-JUL-1990

;; PRIOR APPLICATION DATA:

;; FILING DATE: 24-FEB-1989

;; APPLICATION NUMBER: 315,331

;; FILING DATE: 24-FEB-1989

;; APPLICATION NUMBER: 726,500

;; FILING DATE: 24-APR-1985

;; APPLICATION NUMBER: 489,786

;; FILING DATE: 29-APR-1983

;; SEQ ID NO:1:

;; LENGTH: 152

5196335-1

Query Match

Best Local Similarity 97.6%; Score 801.5; DB 6; Length 152;

Matches 152; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 2 ATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVFHFGDNTAGCTSA 61

Db 1 ATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVFHFGDNTAGCTSA 59

Qy 62 GPHFNPLSRKHGGPKDERHVGDLGNVTADKGVADVSDVSIKGLTEGLHGFHVFHFGDNTAGCTSA 121

Db 60 GPHFNPLSRKHGGPKDERHVGDLGNVTADKGVADVSDVSIKGLTEGLHGFHVFHFGDNTAGCTSA 119

Qy 122 EKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154

Db 120 EKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 152

RESULT 11

US-08-350-884-86

Sequence 86, Application US/08350884

Patent No. 5585258

GENERAL INFORMATION:

APPLICANT: HOUGHTON, MICHAEL

APPLICANT: CHOO, QUI LIM

APPLICANT: KUO, GEORGE

TITLE OF INVENTION: HEPATITIS C VIRUS PROTEASE

NUMBER OF SEQUENCES: 86

CORRESPONDENCE ADDRESS:

ADDRESSEE: MORRISON & FOERSTER

STREET: 755 Page Mill Road

CITY: Palo Alto

STATE: California

COUNTRY: USA

ZIP: 94304-1018

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/350,884

FILING DATE: 06-DEC-1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/680,296

FILING DATE: 04-APR-1991

ATTORNEY/AGENT INFORMATION:

NAME: CIOTTI, THOMAS E.

REGISTRATION NUMBER: 21,013

REFERENCE/DOCKET NUMBER: 22300-20100.20

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 813-5600

TELEFAX: (415) 494-0792

TELEX: 706141

INFORMATION FOR SEQ ID NO: 86:

SEQUENCE CHARACTERISTICS:

LENGTH: 841 amino acids

;; TYPE: amino acid

;; TOPOLOGY: linear

;; MOLECULE TYPE: protein

US-08-350-884-86

Query Match

Best Local Similarity 97.1%; Score 797; DB 1; Length 841;

Matches 149; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVFHFGDNTAGCTSA 60

Db 1 MATNPVCLVKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVFHFGDNTAGCTSA 60

Qy 61 AGPHFNPLSRKHGGPKDERHVGDLGNVTADKGVADVSDVSIKGLTEGLHGFHVFHFGDNTAGCTSA 120

Db 61 PGPHFNPLSRKHGGPKDERHVGDLGNVTADKGVADVSDVSIKGLTEGLHGFHVFHFGDNTAGCTSA 120

Qy 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGI 152

Db 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGI 152

RESULT 12

US-08-709-173-86

Sequence 86, Application US/08709173

Patent No. 5712145

GENERAL INFORMATION:

APPLICANT: HOUGHTON, MICHAEL

APPLICANT: CHOO, QUI LIM

APPLICANT: KUO, GEORGE

TITLE OF INVENTION: HEPATITIS C VIRUS PROTEASE

NUMBER OF SEQUENCES: 86

CORRESPONDENCE ADDRESS:

ADDRESSEE: MORRISON & FOERSTER

STREET: 755 Page Mill Road

CITY: Palo Alto

STATE: California

COUNTRY: USA

ZIP: 94304-1018

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/709,173

FILING DATE: 06-SEP-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/680,296

FILING DATE: 04-APR-1991

ATTORNEY/AGENT INFORMATION:

NAME: CIOTTI, THOMAS E.

REGISTRATION NUMBER: 21,013

REFERENCE/DOCKET NUMBER: 22300-20100.20

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 813-5600

TELEFAX: (415) 494-0792

TELEX: 706141

INFORMATION FOR SEQ ID NO: 86:

SEQUENCE CHARACTERISTICS:

LENGTH: 841 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-709-173-86

Query Match

Best Local Similarity 97.1%; Score 797; DB 1; Length 841;

Matches 149; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVFHFGDNTAGCTSA 60

Db 1 MATNPVCLVKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVFHFGDNTAGCTSA 60

Db 1 MATNPVCVLKGDGPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVEFGDNTAGCTS 60
Qy 61 AGPHFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 120
Db 61 PGPHFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 120
Qy 121 HEKADDLGKGGNEESTKTGNAGSLACGVIGI 152
Db 121 HEKADDLGKGGNEESTKTGNAGSLACGVIGI 152

RESULT 13

US-08-709-177-86
; Sequence 86, Application US/08709177
; Patent No. 5885799
; GENERAL INFORMATION:
; APPLICANT: HOUGHTON, MICHAEL
; APPLICANT: CHOO, QUI LIM
; APPLICANT: KUO, GEORGE
; TITLE OF INVENTION: HEPATITIS C VIRUS PROTEASE
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/709,177
; FILING DATE: 06-SEP-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/680,296
; FILING DATE: 04-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: CIOTTI, THOMAS E.
; REGISTRATION NUMBER: 21,013
; REFERENCE/DOCKET NUMBER: 22300-20100.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 86:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 841 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-709-177-86

Query Match 97.1%; Score 797; DB 2; Length 841;
Best Local Similarity 98.0%; Pred. No. 3.8e-81;
Matches 149; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVEFGDNTAGCTS 60
Db 1 MATNPVCVLKGDGPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVEFGDNTAGCTS 60
Qy 61 AGPHFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 120
Db 61 PGPHFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 120
Qy 121 HEKADDLGKGGNEESTKTGNAGSLACGVIGI 152
Db 121 HEKADDLGKGGNEESTKTGNAGSLACGVIGI 152

RESULT 14
5252476-5
; Patent No. 5252476
; APPLICANT: HALLEWELL, ROBERT A.; MULLENBACH, GUY T.
; TITLE OF INVENTION: SUPEROXIDE DISMUTASE CLONING AND
; EXPRESSION IN MICROORGANISMS
; NUMBER OF SEQUENCES: 15
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/222,352
; FILING DATE: 20-JUL-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 931,920
; FILING DATE: 14-NOV-1986
; APPLICATION NUMBER: 609,412
; FILING DATE: 11-MAY-1984
; APPLICATION NUMBER: 538,607
; FILING DATE: 03-OCT-1983
; SEQ ID NO: 5
; LENGTH: 144
5252476-5

Query Match 94.0%; Score 772; DB 6; Length 144;
Best Local Similarity 100.0%; Pred. No. 2.3e-79;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 11 GDGPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVEFGDNTAGCTSAGPHFNPLSR 70
Db 1 GDGPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVEFGDNTAGCTSAGPHFNPLSR 60
Qy 71 KHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTVVHEKADDLGKG 130
Db 61 KHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTVVHEKADDLGKG 120
Qy 131 GNEESTKTGNAGSLACGVIGIAQ 154
Db 121 GNEESTKTGNAGSLACGVIGIAQ 144

RESULT 15
5171680-3
; Patent No. 5171680
; APPLICANT: MULLENBACH, GUY T.; HALLEWELL, ROBERT A.; VALEZUELA,
; PABLO
; TITLE OF INVENTION: SUPEROXIDE DISMUTASE ANALOGS HAVING NOVEL
; BINDING PROPERTIES
; NUMBER OF SEQUENCES: 15
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/561,442
; FILING DATE: 01-AUG-1990
; SEQ ID NO: 3
; LENGTH: 152
5171680-3

Query Match 92.3%; Score 758; DB 6; Length 152;
Best Local Similarity 94.8%; Pred. No. 9.3e-78;
Matches 146; Conservative 1; Mismatches 5; Indels 2; Gaps 1;
Qy 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVEFGDNTAGCTS 60
Db 1 MATKAVAVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVEFGDNTAGCTS 58
Qy 61 AGPHFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 120
Db 59 AGPHFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 118
Qy 121 HEKADDLGKGGNEESTKTGNAGSLACGVIGIAQ 154
Db 119 HEKADDLGKGGNEESTKTGNAGSLACGVIGIAQ 152

Search completed: January 28, 2003, 14:21:42
Job time : 5.19694 secs

Db 401 ELLAGGSGSPVLSRKQKGVLLGEBEALDDSESRDVSFAAASVKSEIGGELAAS 460
QY 481 SGVSTPGSAGHDITEPRSQHTLOADSVLASCGLTSSATDGEDEEDILSHSSQVSANVP 540
Db 461 SGVSTPGSVGHDITEPRSQHTLOADSVLASCGLTSSATDGEDEEDILSHSSQVSANVP 520
QY 541 SDPAMDNDCTQASSPSSDSSQTTTEGPDASVTPSDSEIVLDTGTDNQVGLGQIGQODE 600
Db 521 PDPPMDLNDCTQSSPSSDSSQTTTEGPDASVTPSDSEIVLDTGTDNQVGLGQIGQODE 580
QY 601 DEE-ATGILDPDEASEAFRNSMALQQAHLKNNSHCRQPSDSDSVDFVLRDEATEPGDQE 659
Db 581 DEEAGAGVLSGEVSDVFRNSLALQOHTLLERMGHSRQPSDSDSIDKYVTRDEVAEASDPE 640
QY 660 NKPCRIGDITGOSTDDSDAPLHVCHVRLLSASFLITGCKNVLPDRDVRVSVKALALSCVG 719
Db 641 SKPCRIGDITGQPNDDSDAPLHVCHVRLLSASFLITGCKNVLPDRDVRVSVKALALSCVG 700
QY 720 AAVLHPESFSLKYKVPDITTEYBEEQVVSIDLNTIDHGDPOVRCATAILCGTLICSTL 779
Db 701 AAVLHPESFSLKYKVPDITTEYBEEQVVSIDLNTIDHGDPOVRCATAILCGTLICSTL 760
QY 780 SRSRHFVDMGIRITLGTNTFSLADCIPLLRKTLKDESSVTCCKLACTAVRNCVMSLCS 839
Db 761 SRSRLVGDWLGIRITLGTNTFSLADCIPLLRKTLKDESSVTCCKLACTAVRNCVMSLCS 820
QY 840 SYSGLGLIIDLITLNRNSYMLVRLTETLAEIDFRLVSLFEAKAENLHGAHHYTG 899
Db 821 SYSGLGLIIDLITLNRNSYMLVRLTETLAEIDFRLVSLFEAKAENLHGAHHYTG 880
QY 900 LKQERVLNNVYIHLGDEDPVRVHVAASLRLVPLKFKYKQDQADPVAVARDQSSV 959
Db 881 LKQERVLNNVYIHLGDEDPVRVHVAASLRLVPLKFKYKQDQADPVAVARDQSSV 940
QY 960 YLKLMLHETQPPSHFVSITRIYRYNLLPSITDVTMNNLSRVTAAVSHELITSTTRA 1019
Db 941 YLKLMLHETQPPSHFVSITRIYRYNLLPSITDVTMNNLSRVTAAVSHELITSTTRA 1000
QY 1020 LTFGCCEALCLLSTAPVPCVWSLGHGCVPPPLSASDESRSKCTVGMATMILTLISSANFP 1079
Db 1001 LTFGCCEALCLLSTAPVPCVWSLGHGCVPPPLSASDESRSKCTVGMATMILTLISSANFP 1060
QY 1080 LDLSAHQDALILAGNLLAASAPKSRSSWASBEEANPAATKOEYVMPALGDRLVPMVEQ 1139
Db 1061 LDLSAHQDALILAGNLLAASAPKSRSSWASBEEANPAATKOEYVMPALGDRLVPMVEQ 1120
QY 1140 LFSHLLKVINICAHVLDVDPGPAKALPSLTNPSPSPIRRKKEKEPEGQASVPLSP 1199
Db 1121 LFSHLLKVINICAHVLDVDPGPAKALPSLTNPSPSPIRRKKEKEPEGQASVPLSP 1180
QY 1200 KKGSEASASROSSTSGPVTTSSKSSLSGSFYHLPVSLKLDHVLKATHANYKVTLDLQNST 1259
Db 1181 KKGSEASASROSSTSGPVTTSSKSSLSGSFYHLPVSLKLDHVLKATHANYKVTLDLQNST 1240
QY 1260 EKFGFLRSALDVLQSOILELATLQIGKCVBEILGLKSCFREPMMATVVCQQLKTLF 1319
Db 1241 EKFGFLRSALDVLQSOILELATLQIGKCVBEILGLKSCFREPMMATVVCQQLKTLF 1300
QY 1320 GTNLASQFDGLSSNPKSQRAQLGSSSVRPLGYHYCFMAYTHFTQALDASLRNMVQ 1379
Db 1301 GTNLASQFDGLSSNPKSQRAQLGSSSVRPLGYHYCFMAYTHFTQALDASLRNMVQ 1360
QY 1380 AQENDTSGWFDVLOKVSQTLKNTITSVTKNRADKNAIHNHRLPEPLVIRKALKQYTTTT 1439
Db 1361 AQENDTSGWFDVLOKVSQTLKNTITSVTKNRADKNAIHNHRLPEPLVIRKALKQYTTTT 1420
QY 1440 CYOLOKQVLDLQAOLVQLRVNVCCLDSQVFTGFLVKOFIEVQCFRESEALIPNIEFF 1499
Db 1421 CYOLOKQVLDLQAOLVQLRVNVCCLDSQVFTGFLVKOFIEVQCFRESEALIPNIEFF 1480
QY 1500 LVLLSYRHSKQIIGIKPIIQLCDGIMASGRKA 1533
Db 1481 LVLLSYRHSKQIIGIKPIIQLCDGIMASGRKA 1514

RESULT 3

T48814

hypothetical protein 15E6.220 [imported] - Neurospora crassa

C:Species: Neurospora crassa

C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 05-May-2000

C:Accession: T48814

R:Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatu

submitted to the Protein Sequence Database, April 2000

A:Reference number: Z24541

A:Accession: T48814

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1952 <SCH>

A:Cross-references: EMBL:AL353822; GSPDB:GN00112; NCSP:15E6.220

A:Experimental source: cosmid contig 15E6; strain 74

C:Genetics:

A:Gene: NCSP:15E6.220

A:Map position: 2

A:Introns: 281/3

Query Match

3.5%; Score 278.5; DB 2; Length 1952;

Best Local Similarity 20.7%; pred. No. 3e-07;

Matches 176; Conservative 79; Mismatches 277; Indels 317; Gaps 31;

QY 18 QQQQQQQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 75

Db 27 QQQQQQQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 80

QY 76 GPAVAEELHPRKELSKATKDRVNHCTICENIVAQSVRNSPEFKLLGTAMELFLICS 135

Db 81 EQPAPPPEPTPS-----QSQPQ-QSL----- 102

QY 136 DDAESDVRWADCELVKIKALMDSNLPRLQLELYKEIKKNGAPRSRAALWRFAELAH 195

Db 103 ---QPQVQLDAPAHKLDLSALENGI-----SYTLNSPPR-----PAISPL 140

QY 196 VRPKCRPYLVNLLPCLTRTSKRPEESVOET-LAAAVPKIMA-SFGNPFANDNEIKVLLKA 253

Db 141 TDVPD-----PRTREPSVQPLESELTNHAADQPKNGIAHSNPTNPNPTDLKGR 192

QY 254 FIANLKSSPTIRTAAGSAVSICQHSRTQYFYSWLLNVLGLLVPEDEHSLTLLIGV 313

Db 193 DTPQVTSVGTAGATVGGAAAT-----TTVPGNSTNSTSHFSG 231

QY 314 LLTLYLVPLQLQQVKDTSLKSGFVTRKEMEVSAPSABQLVQVYELTLHHHTQHODHNVVT 373

Db 232 AMDGTPIAPFNTQAPAVD-----GMANMDF---PSS-----TMDLLQSSGHLIM 274

QY 374 GALELLOQLFRTPPP-----ELLQTLTAVGGIGQLT-----AAK 407

Db 275 SFLOSLSGGIQQPPPPAAAAATAADASLESFARLEFADGVFMOTTYALTIGRDQRAWRLAKK 334

QY 408 EESGGRSRSGSIVELIAGGSSCSFPV---LSRKOKGV-----LLGEEALEDDSESR-- 457

Db 335 EERRAEQYQLKDEYEAQGLTPAPPSEDARRFSKYSISEEGMLGPESDDEDEEGRPS 394

QY 458 -----SDVSSSALTA-SVKDEISGELAASSGV-TPGSAGHDII 494

Db 395 DNRAPKXKRMNGGVSLPVDSEGAESSLMSADQTVDDGKGPVSNRQVSVHTPGAAVNL 454

QY 495 TEQPRSOHT-----NEVPMVSDPDRDSSQSELSEPPDEILLDPADPHVNETV 503

Db 455 ALRSPHHTPLGLTHSPCPNTAAKTKAISREHLKIQPNQAGVFEATPLHKNGFCEBDVH 514

QY 504 LOADSVDLASCD-----LTSSATDG-----DEE 526

Db 515 YSHDKVVLKSGDRQLQVKDVEFVFIINGVABGKTCAEYEPEETPARRYSEGGKEMSDFFE 574

QY 527 DILSHSSQVSANVPDAMDINDGTQASSPISDSSQT-TTEGPDASVTPSDSEIVLDGT 585

Db 575 SI--HDVDRRSTSPED-----NEVPMVSDPDRDSSQSELSEPPDEILLDPADPHVNETV 627

Db 477 SKAVATWANTERQKETERIEKERRMRLMAEDESRYKLID-----QKDRRLA 527
 QY 373 TGALELLOQLFRTPPELLOTLI-AVGGIGQLTAANKESGGRSRSGSIVELIAGGSSCS 431
 Db 528 -----YLLQO-----TDEYVANLVNWEHKQAQAAKEKKRRKKKBAENAEAGGESALG 578
 QY 432 P-----VLSRKQKGVLLGEEBALEDDSESRSDDVSSSALTASVKDEISG 475
 Db 579 PDGEIDESSQMSDLPVKVYTHTTGKVLFGPEAP-----KASOLDRAWLENNPGY 627
 QY 476 ELAASSGVSTPGSAGHDITEQPSRSHTLQADSVOLASCDLTSSATDGEEDILSHSSSQ 535
 Db 628 EVA-----PRS-----DSESDSDYEEDDEEERSRQ 654
 QY 536 -----VSAPSPDAMDNDGTO-----ASSPISD-----SSQTTTECPDSAVTPSDS----- 577
 Db 655 ETEKILLDPNSEVESEKAKIETAKQDVDEYSQMSARGSQSYTTVAHAISERVEK 714
 QY 578 -SEIVLDGTDNOY--LGLQITGQPQDDEBATGILPDE 611
 Db 715 QSALLINGTLKHVQLOGLE-WWVSLYNNNLINGILADE 750

RESULT 6
 AI2043
 hypothetical protein alr1903 [imported] - Nostoc sp. (strain PCC 7120)
 C:Species: Nostoc sp.
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
 C:Accession: AI2043
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriuchika
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
 DNA Res. 8, 205-213, 2001
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
 A:Reference number: AB1807; MUID:21595285; PMID:11759840
 A:Accession: AI2043
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1547 <KUR>
 A:Cross-references: GB:BA000019; PIDN:BA073602.1; PID:g17130993; GSPDB:GN00179
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: alr1903

Query Match 2.8%; Score 220.5; DB 2; Length 1547;
 Best Local Similarity 19.5%; Pred. No. 0.00041;
 Matches 237; Conservative 185; Mismatches 412; Indels 379; Gaps 54;

QY 81 EEPHPRKKELSATKDRVNHCLTICENIVAQSVRNSPEFQKLLGTAMELFLLCSDDAES 140
 Db 355 DEGFNPRDQASRLKDLNHL-----IQLAANDQIEF 388
 QY 141 DVMWAD-----ECLNKKVIALMDSNLPR-----LQLELYEIKKNGAPRSLR 183
 Db 389 RHQIOEYTAECCLLPLSDSDSLQRELYNLKWTPELKLMLDELVD-----R 438
 QY 184 AALWRAELAHVLPQKCRPYLVNLLPCLTRTSKRPEESVQETLAAAVPKIMASFGNFAN 243
 Db 439 ALAVRVNVLAKVDQ-----LGAKLAGAAKPE--LQOQTVDLIPGL----- 478
 QY 244 DNEIKVLLKAFITANKSSPTIRRTAAGSAVSICQHSRRRTQVYFWLLNVLGLLVPVED 303
 Db 479 --KIPLLFKICLLGLTKSEKAISH-----LNPLL-----ED 507
 QY 304 EUSTL-----LILGLV---LTRYLVPLLQOQVKDTLSKSGVGRKMEVSPSAEQVLQ 355
 Db 508 ENSSVRDASAALGEIKSEAAIPGLTIKLEHE--NSSVRYSAASALGEIKSEAAIPGLIN 565
 QY 356 VVELTLHHHTQHODHNVVTGALELLOQL-FRTPPPELLQTL-----TAVGGIGQLTA 405
 Db 566 LLE-----HENFDVRYRAASALGEIKSEAAIPGLINLLEHENS SVRYRAASALGQIKS 618

QY 406 AKEESGGRSRSGSIVELIAGGSS-----CSPVLSRKQKGVLLGEEBALEDDSESRSVDS 461
 Db 619 EQAIIIG-----LILKLEHENS SVRYRAASALGQIKSEAAIPGLTIKLEHEN-----SDVR 668
 QY 462 SSALTASVKDEISGELAASSGV-----TPGSAGHDITTEQ--PRSOHTLOADS 508
 Db 669 YRA--ASALGEIKSEAAIPGLINLLEHENFDVRYSAASALGEIKSEAAIPGLINLLEHEN 726
 QY 509 VDL-----ASCDLTSSAT-----DGEEDILSHSSSOVS AVSPSPAMD--LNDGTQ 552
 Db 727 FDVRYSAASALGEIKSEAAIPGLINLLEHENFDVRYSAASALGEIKSEAAIPGLINLLEH 786
 QY 553 ASSPISDSQTTT-EGPDSAVTPSDSSSIVLDGTDNOYVLGQ-ICQPODED--EATGIL 608
 Db 787 ENFDVRYSAASALGEIKSEAAIPGLINLLEHENFDVRYSAASALGEIKSEAAIPGLINL 846
 QY 609 PDEASEAFNFSMALQO-----AHLKLN-MSHCRQPSDSSVDKRVLRDEATEPG- 656
 Db 847 EDENS DVRYSAASALGEIKSEAAIPGLINLLEHENS PYRRRAASALGO--IKSEAAIPGL 904
 QY 657 ----DOENKPCR-----TKGDIGOSTDDSDAPLVHCVRLLSASF----- 692
 Db 905 INLLEHENS PYRRRAASALGQIKSEAAIPGLINLLEHENS-----VRYSAASSLGQIKS 959
 QY 693 ---LTGGKNVLVPDRDVRVSVKALALSCVGAVALHPESFFSKLYKVPDLTTEPBEQVY 749
 Db 960 EQAIIIGLTKLL--EDENSSVRRRAASALG-----ETKSEQAI 994
 QY 750 SDILNYIDHGPQVGRGATAILCGTILCSILSRSFHVGDMMGTIRLTGNTFSLADCIPL 809
 Db 995 PGLIKLLEHEDSDV-----RYRAASALGEIKS-----EQAI PG 1027
 QY 810 LRKTLKDESSVTCKLACTAVRNCVMSLCSSSYSELGLQIILVTLRNSSVLVRTELLE 869
 Db 1028 LIKLEDEDSE-----VRYRAASALGEIKSEAAIPGLINLLEHEDSS--VRYRAAS 1076
 QY 870 TLAEDIFR-----LVSELEAKAENLRGAAHYTGLLKQERVLNNVTHLLGDEDPVRH 924
 Db 1077 TLGQIKSEAAIPGLINLLEHEDSKSVRDSAAALGEIKSEAAIPG--LINLLEHENS SVRD 1134
 QY 925 VAAASLIRLVKPLFKYCKDQGOADPVAVARDQSSVYLKLLMHETQPPSHFSVSTITRIY 984
 Db 1135 SAASALGEI-----KSEQAIPMLNRLNEEFVSAANNNTLYS 1172
 QY 985 GYNLLPSITDVTMEN-NLSRVIAAVSHELITSTTRALTFCCEALCLLSTAFPCVWSLG 1043
 Db 1173 ALEALEAI-----QENCQYIRCLTEQKQLYLPSPQ-----SKTSLMVLHLS 1217
 QY 1044 WHCGVPPLSASDESRSKCTVGMATMILTLLSSAW---FPDLDS-----AHODALILAGNL 1095
 Db 1218 LHFCTP-----DEANK-----WSNQLAMDCLKQELQIPHLDALILSGDI 1255
 QY 1096 LAASAP-----KSLRSSWASEEENPAATKQEEV-WPALGDRALVPMVQLFSHL 1144
 Db 1256 ANKSTPDEYAAQAQFLDELRODFSLKFEQIIAPGNHDLNMQISEEGYIPTLRKKQES 1315
 QY 1145 LKVINICAHVLDD 1157
 Db 1316 MD--ESC--VIDD 1324

RESULT 7
149681

glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 03-Jun-2002
 C:Accession: I49681; I49046
 C:Welch, J.E.; Schatte, E.C.; O'Brien, D.A.; Eddy, E.M.
 Biol. Reprod. 46, 869-878, 1992
 A:Title: Expression of a glyceraldehyde 3-phosphate dehydrogenase gene specific to mo
 A:Reference number: I49681; MUID:92273722; PMID:1375514
 A:Accession: I49681
 A>Status: preliminary; translated from GB/EMBL/DBJ

Tue Jan 28 14:23:37 2003

us-09-904-987-7.rpr

Page 10

Search completed: January 28, 2003, 14:20:44
Job time : 71.4611 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:03:18 ; Search time 23.6023 Seconds
(without alignments)
2711.515 Million cell updates/sec

Title: US-09-904-987-7

Perfect score: 7892

Sequence: 1 MATLEKLMKAFESLKSQQ.....DGIWASGRKASQPQRLCSP 1543

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match	Length	DB	ID	Description
1	7821.5	99.1	3144	1	HD_HUMAN	P42858	homo sapien
2	7136.5	90.4	3119	1	HD_MOUSE	P42859	mus musculus
3	7045.5	89.3	3110	1	HD_MOUSE	P51111	rattus norv
4	5315	67.3	3148	1	HD_FUGRU	P51112	fugu rubrip
5	235	3.0	875	1	Y066_NPYOP	Q83949	orgyia pseu
6	215.5	2.7	440	1	G3PT_MOUSE	Q64467	mus musculus
7	214	2.7	1586	1	SN22_HUMAN	P51531	homo sapien
8	206	2.6	1822	1	ZAP3_HUMAN	P49750	homo sapien
9	203	2.6	1206	1	FM14_MOUSE	Q05859	mus musculus
10	199.5	2.5	141	1	YPRO_OWEFU	P21260	owenia fusi
11	197	2.5	901	1	Y298_HUMAN	Q15016	homo sapien
12	195.5	2.5	3703	1	ABF1_HUMAN	Q15911	homo sapien
13	194	2.5	1468	1	FMN1_MOUSE	Q05860	mus musculus
14	193.5	2.5	1365	1	SU22_DROME	P25172	drosofila
15	193.5	2.5	3726	1	ABF1_MOUSE	Q61329	mus musculus
16	189	2.4	1213	1	FMN_CHICK	Q05858	gallus gall
17	188	2.4	485	1	SSGP_VOLCA	P21937	volvox cart
18	186	2.4	816	1	HUNB_DROVI	P13361	drosofila
19	185.5	2.4	1790	1	SEPA_EMENI	P78621	emericeila
20	185.5	2.4	3149	1	TEGU_EBV	P03186	epstein-bar
21	183.5	2.3	4385	1	YF73_CABEL	Q09222	caenorhabdi
22	179.5	2.3	1192	1	RTM4_HUMAN	Q9nqc3	homo sapien
23	179	2.3	1362	1	BRD4_HUMAN	Q60885	homo sapien
24	177.5	2.2	1443	1	E75C_DROME	P13055	drosofila
25	177	2.2	5430	1	ACF7_HUMAN	Q9unp3	homo sapien
26	176.5	2.2	487	1	EBN2_EBV	P12978	epstein-bar
27	176.5	2.2	1306	1	MSB2_YEAST	P32334	saccharomyc
28	176	2.2	2004	1	MOZ_HUMAN	Q92794	homo sapien
29	175.5	2.2	1849	1	TI72_HUMAN	O14981	homo sapien
30	174	2.2	3164	1	TEGU_HSV11	P10220	herpes simp
31	173	2.2	1091	1	DIA_DROME	P48608	drosofila
32	171.5	2.2	2670	1	YAO5_SCHPO	O10105	schizosacch
33	171	2.2	905	1	SNF5_YEAST	P18480	saccharomyc

34	170.5	2.2	764	1	SAT1_MOUSE	Q60611	mus musculus
35	170.5	2.2	2715	1	TRX2_HUMAN	Q9unn6	homo sapien
36	170	2.2	797	1	PAT1_YEAST	P25644	saccharomyc
37	168	2.1	1217	1	AF4_MOUSE	O88573	mus musculus
38	167.5	2.1	497	1	WAS2_HUMAN	Q9y6w3	homo sapien
39	167	2.1	1902	1	SMF1_HUMAN	O14497	homo sapien
40	166.5	2.1	853	1	DIA3_HUMAN	Q9nsv4	homo sapien
41	166	2.1	1140	1	YM96_YEAST	Q04893	saccharomyc
42	166	2.1	1171	1	DIA3_MOUSE	Q92207	mus musculus
43	163	2.1	431	1	ACRO_RABIT	P48038	oryctolagus
44	163	2.1	1059	1	CAPU_DROME	Q24120	drosofila
45	162.5	2.1	2426	1	SON_HUMAN	P18583	homo sapien

ALIGNMENTS

RESULT 1
HD_HUMAN
ID HD_HUMAN STANDARD; PRT; 3144 AA.
AC P42858;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Huntingtin (Huntington's disease protein) (HD protein).
GN HD OR IT15.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Retina;
RX MEDLINE=93208892; PubMed=8458085;
RA MacDonald M., Ambrose C.M., Duyao M.P., Myers R.H., Lin C.S.,
RA Jenkins J., Barnes G., Taylor S.A., James M., Groot N., McFarlane H.,
RA Strindhi J., Anderson M.A., Wexler N.S., Gusella J.F., Bates G.P.,
RA Baxendale S., Hummerich H., Kirby S., North M., Youngman S., Mott R.,
RA Zehner G., Sedlacek Z., Poukette A., Frischau A.-M., Lehrach H.,
RA Buckler A.J., Church D., Doucette-Stamm L., O'Donovan M.C.,
RA Riba-Ramirez L., Shah M., Stanton V.P., Strobel S.A., Draths K.M.,
RA Wales J.L., Dervan P., Housman D.E., Altherr M., Shiang R.,
RA Thompson L., Fielder T., Wasmuth J.J., Tagle D., Valdes J., Elmer L.,
RA Allard M., Castilla L., Swarcop M., Blanchard K., Collins F.S.,
RA Snell R., Holloway T., Gillespie K., Datson N., Shaw S., Harper P.S.;
RT "A novel gene containing a trinucleotide repeat that is expanded and
RT unstable on Huntington's disease chromosomes. The Huntington's
RT Disease Collaborative Research Group.";
RL Cell 72:971-983(1993).
RN [2]
RP SEQUENCE OF 1-90 FROM N.A.
RX MEDLINE=95278941; PubMed=7759106;
RA Lin B., Nasir J., Kaichan M.A., McDonald H., Zeisler J.,
RA Goldberg Y.P., Hayden M.R.;
RT "Structural analysis of the 5' region of mouse and human Huntingtin
RT di- and trinucleotide polymorphisms.";
RL Genomics 25:707-715(1995).
RN [3]
RP SEQUENCE OF 1-205 FROM N.A.
RX MEDLINE=94255787; PubMed=8197474;
RA Ambrose C.M., Duyao M.P., Barnes G., Bates G.P., Lin C.S.,
RA Strindhi J., Baxendale S., Hummerich H., Lehrach H., Altherr M.,
RA Wasmuth J., Buckler A., Church D., Housman D., Berks M., Micklem G.,
RA Durbin R., Dodge A., Read A., Gusella J.F., Macdonald M.E.;
RT "Structure and expression of the Huntington's disease gene: evidence
RT against simple inactivation due to an expanded CAG repeat.";
RL Somatic. Cell Mol. Genet. 20:27-38(1994).
RN [4]
RP SEQUENCE OF 1-117 FROM N.A.
RA Matthews P.;
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
RN [5]

QY 298 LVPVEDEHSTLLILGVLRLYLVPLLOQVKDTSKSGFVTRKEMEYSPSAEQLVQY 357
 Db 301 LVPVEDEHSTLLILGVLRLYLVPLLOQVKDTSKSGFVTRKEMEYSPSAEQLVQY 360
 QY 358 ELTLHTHQHNDVNTGALQLQLFRTPPPELLQTLTAVGGIGQLTAAKEESGGRSRG 417
 Db 361 ELTLHTHQHNDVNTGALQLQLFRTPPPELLQTLTAVGGIGQLTAAKEESGGRSRG 420
 QY 418 SIVELLAGGSGSPVLSRKQKGVLLGEEALEDDSESRDVSASSALTASVKDEISGEL 477
 Db 421 SIVELLAGGSGSPVLSRKQKGVLLGEEALEDDSESRDVSASSALTASVKDEISGEL 480
 QY 478 AASSGVSTPCSGAGHDIITEQPSQHTLOQDSVDLASCDLTSSATDGEDDILSHSSQVS 537
 Db 481 AASSGVSTPCSGAGHDIITEQPSQHTLOQDSVDLASCDLTSSATDGEDDILSHSSQVS 540
 QY 538 AVPSDPMNDNGTQASSPISDSSTTEGPDPSAVTPSDSSIVLDGTDNQYGLQIGQP 597
 Db 541 AVPSDPMNDNGTQASSPISDSSTTEGPDPSAVTPSDSSIVLDGTDNQYGLQIGQP 600
 QY 598 QDEDEATGILPDEASEAFRNSMALQQAHLKNNKSHCHROPSDSSVDKFLVLRDEATEPGD 657
 Db 601 QDEDEATGILPDEASEAFRNSMALQQAHLKNNKSHCHROPSDSSVDKFLVLRDEATEPGD 660
 QY 658 QENKPCRIGDITQSDDDSDAPLVHCVRLLSASFLLTGKKNVLPDRDVRVSVKALALSC 717
 Db 661 QENKPCRIGDITQSDDDSDAPLVHCVRLLSASFLLTGKKNVLPDRDVRVSVKALALSC 720
 QY 718 VCAAVALLHPESFSLKYKVPDLTTEPEQYVSDILNTIDHGDPOVRGATAILCGTLICS 777
 Db 721 VCAAVALLHPESFSLKYKVPDLTTEPEQYVSDILNTIDHGDPOVRGATAILCGTLICS 780
 QY 778 ILSRSRHFVGDWMTGIRTLTGNTFSLADICPLLRKTLKDESSVTKCLACTAVRNCVMSLC 837
 Db 781 ILSRSRHFVGDWMTGIRTLTGNTFSLADICPLLRKTLKDESSVTKCLACTAVRNCVMSLC 840
 QY 838 SSSYSELGLQLIIDVLTRNSSYWLVRTELLETAEIDFRLVFLFLEAKAENLHRGAHYT 897
 Db 841 SSSYSELGLQLIIDVLTRNSSYWLVRTELLETAEIDFRLVFLFLEAKAENLHRGAHYT 900
 QY 898 GLKQLQERVNNVTHLLGDEDPVRHVAAASLRLVPLKFLYKCDQGDQADPVAVARDQS 957
 Db 901 GLKQLQERVNNVTHLLGDEDPVRHVAAASLRLVPLKFLYKCDQGDQADPVAVARDQS 960
 QY 958 SVYLLKLMHETQPPSHFSVITIRYGYNLLPSTDTVTMNNLSRVIAAASHLITSTT 1017
 Db 961 SVYLLKLMHETQPPSHFSVITIRYGYNLLPSTDTVTMNNLSRVIAAASHLITSTT 1020
 QY 1018 RALTFGCCCEALCLLSTAPPVCIVSLGWCHGVPPLISASDESRSKCTVGMATMILTLLSSAW 1077
 Db 1021 RALTFGCCCEALCLLSTAPPVCIVSLGWCHGVPPLISASDESRSKCTVGMATMILTLLSSAW 1080
 QY 1078 FPLDLSAHODALILAGNLLAASAPKSLRSSWASSEENAPATKQEEVWPALGDRALVPMV 1137
 Db 1081 FPLDLSAHODALILAGNLLAASAPKSLRSSWASSEENAPATKQEEVWPALGDRALVPMV 1140
 QY 1138 EQLFSLHLKLVINICAHVLDVAVGPAIKAAPLSLTNPSPISPIRRKKEKEPGEQASVPL 1197
 Db 1141 EQLFSLHLKLVINICAHVLDVAVGPAIKAAPLSLTNPSPISPIRRKKEKEPGEQASVPL 1200
 QY 1198 SPKKGSEASASRSQSDTSGPVTTSKSSSLGSGFYHLPSYLKLDHVKATHANYKVTLDLQN 1257
 Db 1201 SPKKGSEASASRSQSDTSGPVTTSKSSSLGSGFYHLPSYLKLDHVKATHANYKVTLDLQN 1260
 QY 1258 STEKFGFLRSALDVLQILELATLQDICKVEELGVLKSCFREPDMATVCVQQLLKT 1317
 Db 1261 STEKFGFLRSALDVLQILELATLQDICKVEELGVLKSCFREPDMATVCVQQLLKT 1320
 QY 1318 LFGTNLASQFDGLSNFSPKSGRAQLGSSVRPGLYHYCFMAYPTHFTQALADASLRNM 1377
 Db 1321 LFGTNLASQFDGLSNFSPKSGRAQLGSSVRPGLYHYCFMAYPTHFTQALADASLRNM 1380
 QY 1378 VQAEQENDTSGWFDVLOKVSTQKNTLTSTVTKNRADKNAIHNHRLFEPLVIKALKOYTT 1437

Db 1381 VQAEQENDTSGWFDVLOKVSTQKNTLTSTVTKNRADKNAIHNHRLFEPLVIKALKOYTT 1440
 QY 1438 TTCVLOKQVLDLALQVLRVNYCLDSDOVFGVLKQFEYIEVGFQFSEAIIPNIF 1497
 Db 1441 TTCVLOKQVLDLALQVLRVNYCLDSDOVFGVLKQFEYIEVGFQFSEAIIPNIF 1500
 QY 1498 FFLVLLSYERYHSHKOITIGPKIIQLCDGIMASGRKA 1533
 Db 1501 FFLVLLSYERYHSHKOITIGPKIIQLCDGIMASGRKA 1536
 RESULT 2
 HD_MOUSE HD_MOUSE STANDARD; PRT; 3119 AA.
 AC P42859;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Huntingtin (Huntington's disease protein homolog) (HD protein).
 GN HD OR HDH.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS).
 RC STRAIN=C57BL/6; TISSUE=Brain, and Spleen;
 RX MEDLINE=94214482; PubMed=8162057;
 RA Lin B., Nasir J., Macdonald H., Hutchinson G., Graham R.K.,
 RA Rommens J.M., Hayden M.R.;
 RT "Sequence of the murine Huntington disease gene: evidence for
 RT conservation, alternate splicing and polymorphism in a triplet (CCG
 RT repeat.";
 RL Hum. Mol. Genet. 3:85-92(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94278649; PubMed=80093370;
 RA Barnes G.T., Duyao M.P., Ambrose C.M., McNeill S., Persichetti F.,
 RA Srinidhi J., Gusella J.F., Macdonald M.E.;
 RT "Mouse Huntington's disease gene homolog (Hdh).";
 RL Somat. Cell Mol. Genet. 20:87-97(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95375771; PubMed=7647777;
 RA Trotter Y., Devys D., Imbert G., Saudou F., An I., Lutz Y.,
 RA Weber C., Agid Y., Hirsch E.C., Mandel J.L.;
 RT "Cellular localization of the Huntington's disease protein and
 RT discrimination of the normal and mutated form.";
 RL Nat. Genet. 10:104-110(1995).
 RN [4]
 RP SEQUENCE OF 1-181 FROM N.A.
 RX MEDLINE=95278941; PubMed=7759106;
 RA Lin B., Nasir J., Kaichman M.A., McDonald H., Zeisler J.,
 RA Goldberg Y.P., Hayden M.R.;
 RT "Structural analysis of the 5' region of mouse and human Huntington
 RT disease genes reveals conservation of putative promoter region and
 RT di- and trinucleotide polymorphisms.";
 RL Genomics 25:707-715(1995).
 CC -1- FUNCTION: MAY PLAY A ROLE IN MICROTUBULE-MEDIATED TRANSPORT OR
 CC VESICLE FUNCTION.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; a long form (shown here) and a
 CC short form; seem to be produced by alternative splicing. The
 CC short form cannot be explained by a simple splicing event.
 CC -1- TISSUE SPECIFICITY: THE HIGHEST LEVEL IS SEEN THROUGHOUT THE
 CC BRAIN, BUT IT IS ALSO FOUND IN THE STOMACH, HEART, TESTIS, ADIPOSE
 CC TISSUE, MUSCLE, SPLEEN, LIVER, AND KIDNEY.
 CC -1- DEVELOPMENTAL STAGE: PREDOMINANT EXPRESSION IN NEURONAL TISSUES AT
 CC ALL DEVELOPMENTAL STAGES. IN 14.5 DAY OLD EMBRYOS, IT IS ALSO
 CC DETECTED IN NON-NEURONAL TISSUES. THIS EXPRESSION IS DOWN-
 CC REGULATED IN LATER STAGES OF DEVELOPMENT.
 CC -1- POLYMORPHISM: THE FIRST POLY-PRO REPEAT STRETCH DIFFERS IN LENGTH

Db 106 RE-----LTVINGEGR-----IG-----RLVLRV 125
 QY 149 CLNKVIK--ALMDSNL-PRLOLELYKEIKNGAPSLRAALWRFALHLVPRQKCRPYL 205
 Db 126 CMKGRVAVNDPFDPEYWMYFKYDTHGRYK-----GNVEH-----KNGOLV 171
 QY 206 VNLPLCLTRTSRPEESVOETLAAAVPKIMASFGN--FANDNEIKVLKAFIANLKSSSP 263
 Db 172 VDNLEINTYCKDPKE-----IP--WSSIGNPVVVECTGVYLSSTEASAHSISGAR 220
 QY 264 TIRTAAGSAVSCQHSRRTQFYFWSLLNVLGLLPVVEDEHSTLLILGLVLLTLYLVL 323
 Db 221 RVVVTA-----PSPD--APMFVNGV----- 238
 QY 324 LOOOVRDTSILKSGFVTRKREMEVSPSAQLVQVYELT-----LHHTQHDHNVVTGALE 377
 Db 239 -----NEKDYNGSWTIVSNASCCTNCLAPLAKVITHENFIVEGILM- 279
 QY 378 LLOQLRTPPPPELLOTLTAVGIGOLTAKEESGGRSGSIVELIAGGSGSCPVLRSK 437
 Db 280 -----TTVHSYTATQKTVDG-----PSKDWRGGRGAHQNIIPSSGCAAKAVGVIP-E 327
 QY 438 QKGKVLGEEAELEDDSESRSDV-----SSSALTASVDEISGELA 478
 Db 328 LKKG-LTGMFAVRVTPNVSWDLTCLRLAKPASYSAITAVKAAKGPLA 375

RESULT 7
 SN22.HUMAN
 ID SN22.HUMAN STANDARD; PRT; 1586 AA.
 AC P51531;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Possible global transcription activator SNF2L2 (SNF2-alpha).
 GN SMARCA2 OR SNF2L2 OR BRM OR SNF2A.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=94038910; PubMed=8223438;
 RA Murchard C., Yaniv M.;
 RT "A human homologue of Saccharomyces cerevisiae SNF2/SWI2 and
 RT Drosophila brm genes potentiates transcriptional activation by the
 RT glucocorticoid receptor.";
 RL EMBO J. 12:4279-4290(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=94268902; PubMed=8208605;
 RA Chiba H., Muramatsu M., Nomoto A., Kato H.;
 RT "Two human homologues of Saccharomyces cerevisiae SWI2/SNF2 and
 RT Drosophila brhma are transcriptional coactivators cooperating with
 RT the estrogen receptor and the retinoic acid receptor.";
 RL Nucleic Acids Res. 22:1815-1820(1994).
 CC -!- FUNCTION: TRANSCRIPTIONAL COACTIVATOR COOPERATING WITH NUCLEAR
 CC HORMONE RECEPTORS TO POTENTIATE TRANSCRIPTIONAL ACTIVATION.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- SIMILARITY: CONTAINS 1 BROMODOMAIN.
 CC -!- SIMILARITY: BELONGS TO THE SNF2/RAD54 HELICASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X72889; CAA51407.1; -.

DR EMBL; D26155; BAA05142.1; -.
 DR Genew; HGNC:11098; SMARCA2.
 DR MIM; 600014; -.
 DR InterPro: IPR001487; Bromodomain.
 DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR001650; Helicase_C.
 DR InterPro: IPR000330; SNF2_N.
 DR Pfam; PF00176; SNF2_N; 1.
 DR Pfam; PF00271; Helicase_C; 1.
 DR Pfam; PF00439; bromodomain; 1.
 DR PRINTS; PR00503; BROMODOMAIN.
 DR SMART; SM00297; BROMO; 1.
 DR SMART; SM00487; DEXDC; 1.
 DR SMART; SM00490; HELIC; 1.
 DR PROSITE; PS00633; BROMODOMAIN_1; 1.
 DR PROSITE; PS0014; BROMODOMAIN_2; 1.
 KW Transcription regulation; Nuclear protein; Activator; Bromodomain;
 KW ATP-binding; Helicase; Alternative splicing.
 FT DOMAIN 216 238 POLY-GLN.
 FT DOMAIN 241 249 POLY-GLN.
 FT DOMAIN 555 558 POLY-ARG.
 FT DOMAIN 639 646 POLY-GLU.
 FT NP_BIND 745 752 ATP (POTENTIAL).
 FT SITE 847 850 DEGH BOX.
 FT DOMAIN 1293 1297 POLY-GLU.
 FT DOMAIN 1415 1485 BROMODOMAIN.
 FT DOMAIN 1514 1525 POLY-GLU.
 FT VARSPIC 1397 1414 MISSING (IN SHORT ISOFORM).
 FT CONFLICT 239 239 P -> PQQP (IN REF. 2).
 FT CONFLICT 390 390 Q -> E (IN REF. 2).
 FT CONFLICT 509 509 G -> S (IN REF. 2).
 FT CONFLICT 707 707 W -> R (IN REF. 2).
 FT CONFLICT 1135 1135 D -> H (IN REF. 2).
 FT CONFLICT 1390 1390 C -> V (IN REF. 2).
 SQ SEQUENCE 1586 AA; 180762 MW; FA537E2A2392807A CRC64;

Query Match 2.7%; Score 214; DB 1; Length 1586;
 Best Local Similarity 22.8%; Pred. No. 0.0023;
 Matches 154; Conservative 68; Mismatches 244; Indels 208; Gaps 29;

QY 9 KAFESLKSFOQQQQQQQQQQQQQQQQQQQQPPPPPP-----ppppppolpop- 53
 Db 210 RTPLGLQQQQQQQQQQQQQQQQQQQQPPQPTQQQQQPALVYNYNRPSPGPELSCPS 269
 QY 54 PQQAQPL-----LQOPQPPPPPPPPPP-----PGPAVAEPLHRPKKELSATKKDRVNHCLTI 105
 Db 270 TPQKLPVPAPGGRSPAPPAAPAAAPGSPVPQAPGQPSVLOLQK----- 320
 QY 106 CENIVAQSVRNSPFQKLLGIAMELEFLICSDDAESDVRWVADECLNKVIRKALMDSNLPRL 165
 Db 321 -----QS-RISP-IQPGGL-----DPVE-----ILQEREYRL 346
 QY 166 QLELYKEKK-----NGAPRSLRAALWRFALHLVPRQKCRPYLVNLLPCLTRTSKRPE 220
 Db 347 QARIAHRIQELNLPGLSPDPLRTKATVELKALLNFQ--RLRQEVVACMRDPT----- 400
 QY 221 ESVQETLAAAVPKIMASFGNFANDNEIKVLKAFIANLKSSSTIRTAAGSAVSCQHS 280
 Db 401 -----TLETANS-----KAYKRKQTLREARMTKEKLEKQKIEQE 437
 QY 281 RRTQYFYSWLLNVLGLLPVVEDEHSTLLILGLVLLTLYLVLPL-----QQQVKDTSKLG 335
 Db 438 KRRQKHQOYLNSILQHAQDKKEYHRS--VAGTKOKLSKAVATWANTEREQKETERIE 495
 QY 336 SFGVTRKREMEVSPSAQLVQVYELTLLHHTQHDHNVVTGALELLQLQLFRTPPPELOTLT 395
 Db 496 KERMRRLMAEDEGYRKLID-----QKDRRLA-----YLLQ-----TDEYVANLT 537
 QY 396 -AVGIGIGOLTAAKESGGRSGSIVELIAGGSGSCP-----VLSRQK 438
 Db 538 NLVWEHQAAQAQAKKERRRRKKAENAGGESALGPDGEPIDESSQMSDLPVKVHTFE 597
 QY 439 KGKVLGEEAELEDDSESRSDVSSSALTASVKSDEISGELAASSGVSTPGSAGHDITTEQP 498

SQ	SEQUENCE	1822 AA;	204947 MW;	8B6CB83FE540C7D2 CRC64;
	Query Match	2.6%;	Score 206;	DB 1; Length 1822;
	Best Local Similarity	20.1%;	Pred. No. 0.0073;	
	Matches 159;	Conservative	93;	Mismatches 310; Indels 228; Gaps 29;
QY	16	SFOQQQQQQQQQQQQQQQQQQQQPPPPPPPQLPQPQAQPLLQPQPpppppppp 75		
DB	52	SFREQHLAQLOQLQOMHOKQMOCVQLPHLLPppLPpppVMPGGYGWDQpppppppppp 111		
QY	76	GPAVAEPLHRPKELSGATKKDRYNHCLTTCENIVAOSVRNSPEFQKLLGIAMELFLCLS 135		
DB	112	GPALSYGQQOYKHMLHHQRDP-----PGLVPMELESPP-----S 149		
QY	136	DDAESDVRMVADECLNKVIKALM-----DSNLPRLQELYKEIKKGAPRSLRAALWRF 189		
DB	150	PPVPPGSIYMPQS-----SYMPPPPPSYPPTSSQYLPPLPAQSPSQSPQSYLEA 202		
QY	190	AELAHLVRPOKCRPYL---VNLLPCLRTSKRPESVOETLAAPVKIMAFGNFANDNE 246		
DB	203	PTPSYSSSSSSQSYLSHSQSYLPSSQASPRSQGHKSQALLAPPSPAPPG-----NK 257		
QY	247	IKVLLKAFIANLKSSPTIRRTAGSAVSICQHSHRRITQFYFWELLNLGLLVPEDEHS 306		
DB	258	TTVQOEPLESGAKNKSTEQQAAPEDPDSTWTPEQQQYWYR----- 299		
QY	307	TLILGLVLLTRYLVPLLQQQ-----VKDTSLKSGFGVT---RKMEVSPSA 350		
DB	300	-----QHLLSLQORTKVHLPGHKHGPPVAKDTPPEVKBEIVPATSVQSPESPSS 348		
QY	351	EQ-----LVQVYELTLHTHQDHNV----- 371		
DB	349	EEPLPNNPEEVPPLPPEEPQSEDEADARKLQAAAAHWQHQQHURVGFTQGIMQK 408		
QY	372	---VTGALELLQLFRTPPELLQTLTAVG-----GI-QOLTAAKESEGGRSRSGSI 419		
DB	409	HTQIQIILQOYQIIQ-PPPHIQATTTPPGIPPGVPGQIPPQLTAAPVPPASSSSQSQV 467		
QY	420	VEL-----IAGG---SCSPVLRSKQKGVKLGEAELEDSDSERSDVSSSALTA 467		
DB	468	PEKRPALLTPVFGSAPPTYHPPL-----QSAGPSQVNSKAPLSKALSALPY 516		
QY	468	SVKDEISCELASGSVTFPGSAGHD-----IITEQPRSQHTLQADS-----VD 510		
DB	517	SFSFDQGLGESSAAPSPITAVKDMPVRSGLLPDPRESSY-LESPRGFDPGRREFD 575		
QY	511	LAS-CD-----LTSSATGDGEDILSH-----SSSOVASVPDSPAMD----- 547		
DB	576	LGSRCGRPCKGRFEGNRPDGPRRYEGHPAECTKSWGMGIMPGRPASQFYITPSTLSLP 635		
QY	548	-NDGTQASPISDSSOTTTEGPDSAVTP-SDSSETVLBDGTNOYLGLQIGQODEDEAT 605		
DB	636	RQSGPWKGPATGCQQHQQPKSQAEPLSGNKEPLATSSNQKNFKM-----QSAAF 689		
QY	606	GILPD-----EASEAFRNSSMALQAAHLK-----NMSHCRPSPSSVDVKFVLROEA 652		
DB	690	STAADVDKVKAQSNLENLSDSQEPPKSEVSEGPVPSNWQDNQVQSMETQIDK---AAQV 746		
QY	653	TEPGDQENKP 662		
DB	747	TQPVLANPK 756		
<hr/>				
RESULT 9				
FM14_MOUSE				
ID	FM14_MOUSE	STANDARD;	PRT;	1206 AA.
AC	Q03659;			
DT	01-JUN-1994 (Rel. 29, Created)			
DT	01-JUN-1994 (Rel. 29, Last sequence update)			
DT	30-MAY-2000 (Rel. 39, Last annotation update)			
DE	Formin 1 isoform IV (Limb deformity protein).			
CN	FMN OR LD			
OS	Mus musculus (Mouse).			

Result No.	Score	Query %			DB	ID	Description
		Match	Length				
1	7811.5	99.0	3144	4	Q3U0B7	Q9uqb7 homo sapien	
2	7041	89.2	3139	6	Q9GM99	Q9gm99 sus scrofa	
3	5459.5	69.2	3121	13	Q42369	O42269 brachydanic	
4	674.5	8.5	3563	5	Q9V3N4	Q9v3n4 drosophila	
5	668	8.5	3584	5	Q9Q939	Q9q939 drosophila	
6	358	4.5	75	6	O18897	O18897 canis famila	
7	356.5	4.5	620	5	Q8T069	Q8t069 drosophila	
8	294	3.7	1992	3	Q9P6T1	Q9p6t1 neurospora	
9	293	3.7	67	11	Q9Z1I8	Q9z1i8 rattus norv	
10	229.5	2.9	729	12	Q91TW1	Q91tw1 tupiaa herp	
11	223	2.8	1214	4	Q9B0Q4	Q9b0q4 homo sapien	
12	220.5	2.8	1547	16	Q8VVS1	Q8yvs1 anabaena sp	
13	210	2.7	192	4	Q9UH11	Q9uh11 homo sapien	
14	210	2.7	5038	11	Q9QXX7	Q9qyx7 mus musculu	
15	208	2.6	197	4	Q9UKC2	Q9ukc2 homo sapien	
16	205	2.6	3536	5	Q9VZ30	Q9vz30 drosophila	

Db 114 QKLGIAEMLFLLCSDDAESDVRWVADECLNKVIRKALMDSNLPRLQLELYKEIKNGAPR 173
QY 181 SLRAALWRRFAELAHVLPKCRPYLVNLLPCLTRTSKRPEESVQETLAAAVPKIMASFCN 240
Db 174 SLRAALWRRFAELAHVLPKCRPYLVNLLPCLTRTSKRPEESVQETLAAAVPKIMASFCN 233
QY 241 FANDNEIKVLLKAFIAKLSSTPIRRTAAGSAVSTICQHSRTOFYFYSWLLNLLGLLVP 300
Db 234 FANDNEIKVLLKAFIAKLSSTPIRRTAAGSAVSTICQHSRTOFYFYSWLLNLLGLLVP 293
QY 301 VEDEHSTLLILGVLTLRLVYLLPQQVKTSLKSGFGVTRKEMEVSPSABQLVQVYELT 360
Db 294 VEDEHSTLLILGVLTLRLVYLLPQQVKTSLKSGFGVTRKEMEVSPSABQLVQVYELT 353
QY 361 LHHTQHODHNVVTGALLELQQLFRTPPPPELLQTLTAVGIGOLTAAKPEESGGRSRGSIV 420
Db 354 LHHTQHODHNVVTGALLELQQLFRTPPPPELLQTLTAVGIGOLTAAKPEESGGRSRGSIV 413
QY 421 ELIAGGSSCSPVLRSKQKGVLLGEEBALEDDSDRSVSSALTSVSKDEISGELAAS 480
Db 414 ELIAGGSSCSPVLRSKQKGVLLGEEBALEDDSDRSVSSALTSVSKDEISGELAAS 473
QY 481 SGVSTPGSA-----GHDIITEQPRSOHTLQADSVDLASCDLTSSATDGEDILSHSS 533
Db 474 SGVSTPGSAADSVGHDIITEQPRSOHTLQADSVDLASCDLTSSATDGEDILSHSS 533
QY 534 SQVSAVSPDAMDNDGTQASPSIDSSQTTTEGPDSDSAVTPSDSSEIVLGDTONOYLGLQ 593
Db 534 SQVSAVSPDAMDNDGTQASPSIDSSQTTTEGPDSDSAVTPSDSSEIVLGDTONOYLGLQ 593
QY 594 IQOPO--DEDEATGILPDEASEAFRNSMALQQAHLKNNHCHQPSDSVSKFVLRDE 651
Db 594 LQPOQDADADEDAAGLLPHGGPDAPFRSSPTALQSHVLKSLGHRSQPSDSVSKFVLRDE 653
QY 652 ATEPGDQENKPCRIGKIDTQSTDDSDAPLVHCVRLLSAELLTGGKVLVDPDRVRSVK 711
Db 654 AAEAGEPENKPCRIGKIDTQSTDDSDAPLVHCVRLLSAELLTGGKVLVDPDRVRSVK 713
QY 712 ALALSCVGAVALHPESFSLKLYVLPDTTEYPEEQVSDILNYIDHGDPOVRGATAILC 771
Db 714 ALALSCVGAVALHPESFSLKLYKAPLDSVEYDPEQVSDILSYIHGGDPQVRGATAILC 773
QY 772 GTLICSILSRHFVHGMGTIRTLTGNTPSLADICPLLRKTLKDESSVTKLACTAVRN 831
Db 774 GTLVSSILSRHFVHGMGTIRTLTGNTPSLADICPLLRKTLKDESSVTKLACTAVRN 833
QY 832 CVMSCSSSYSELGLQILIDVTLTRNSSYWLVRTELLETAEIDFRLVSPLEAKAENLHR 891
Db 834 CVAALCGSSYSQWGLQTLTDLTLRSSSYWLVRTELLETVAEIDFRLVSPLEAKAENLHR 893
QY 892 GAHHYTGLLKQERVLNNVYHLLGDEDDPRVHVAASLRLVPLKLYKCDQDQADPVVA 951
Db 894 GAHHYTGLLKQERVLNNVYHLLGDEDDPRVHVAASLRLVPLKLYKCDQDQADPVVA 953
QY 952 VARDQSSVYLKLLMHETOPPSHFVSITRIYRGYNLPLSPITDVTWENLRSVIAVSHS 1011
Db 954 VARDQSSVYLKLLMHETOPPSHFVSITRIYRGYNLPLSPITDVTWENLRSVIAVSHS 1013
QY 1012 LITSTTRALTFCCEALCLLSTAFFVCVWSLGHGCVPLPLSADSRKSCVTVMATMILT 1071
Db 1014 LITSTTRALTFCCEALCLLSTAFFVCVWSLGHGCVPLPLSADSRKSCVTVMATMILT 1073
QY 1072 LLSSAWPPLDLASAHODALILAGNLLAASAPKSLRSWASEEENANPAATKQEEVFPALGDR 1131
Db 1074 LLSSAWPPLDLASAHODALILAGNLLAASAPKSLRSWASEEENANPAATKQEEVFPALGDR 1133
QY 1132 ALVPMVQBLFSLHLLKVLINICAHVLDVADGPAKALPILSTNPPSLPRTKRKEKEPEGE 1191
Db 1134 SLVPMVQBLFSLHLLKVLINICAHVLDVADGPAKALPILSTNPPSLPRTKRKEKEPEGE 1193
QY 1192 QASVPLSPKKGSGSASASROSDTSGPVTTSSSLGSSFYHLPSYKLHDLVAKATHANYKV 1251
Db 1194 QASVPLSPKKGSGSASASROSDTSGPVTTSSSLGSSFYHLPSYKLHDLVAKATHANYKV 1253

QY 1252 TLDLQNSTEFGFLRSALDVLQSLLELATLQDICKVCVEEILGYLKSCFSREPMATVCV 1311
Db 1254 TLDLQNSTEFGFLRSALDVLQSLLELATLQDICKVCVEEILGYLKSCFSREPMATVCV 1313
QY 1312 QQLLKTFLGTNLASQDGLSSNPSQSGRAQRLGSSSVRPGLYHYCFNAPYTHFTQALAD 1371
Db 1314 QQLLKTFLGTNLASQDGLSASTPSKCGRAQRLGSSSVRPGLYHYCFNAPYTHFTQALAD 1373
QY 1372 ASLRNVVQAEOENDTSGWFDVLQKYSTQLKTNLTSTVTKNRADKNAIHHIRLFEPLVKA 1431
Db 1374 ASLRNVVQAEOENDTSGWFDVLQKYSTQLKTNLTSTVTKNRADKNAIHHIRLFEPLVKA 1433
QY 1432 LKQYTTTTTCVQLQKQVLDLLAQLVQLRVNYCLLDSQVDFIGFVLKQFYIEVGOFRESEA 1491
Db 1434 LKQYTTTTTCVQLQKQVLDLLAQLVQLRVNYCLLDSQVDFIGFVLKQFYIEVGOFRESEA 1493
QY 1492 IIPNIFFFVLVLSYERYHSKQIIGIPKIIQICDGMASGRKA 1533
Db 1494 IIPNIFFFVLVLSYERYHSKQIIGIPKIIQICDGMASGRKA 1535

RESULT 3
O42269
ID O42269 PRELIMINARY: PRT: 3121 AA.
AC O42269;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE HUNTINGTIN.
GN HD OR HD.
OS Brachydanio rerio (zebrafish) (zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Karlovich C.A., Ramirez L., John R., Stainier D., Myers R.M.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 9-138 FROM N.A.
RA Williams L.C.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF052602; AAC63983.1; -;
DR EMBL; AF052603; AAC63984.1; -;
DR EMBL; AF028731; AAC05309.1; -;
DR ZFIN; ZDB-GENE-990415-131; hd.
DR InterPro; IPR000091; Huntingtin.
DR Pfam; PF03541; Huntingtin.
DR PRINTS; PR00375; HUNTINGTIN.
SQ SEQUENCE 3121 AA; 346666 MW; 3CC884E199D324AC CRC64;

Query Match 69.2%; Score 5459.5; DB 13; Length 3121;
Best Local Similarity 68.6%; Pred. No. 0;
Matches 1075; Conservative 189; Mismatches 210; Indels 93; Gaps 11;

QY 1 MATLEKLMKAFESLKSTQQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPPPPPPPPPPP 60
Db 1 MATMEKLMKAFESLKSPQQQ----- 21
QY 61 LPQPQPP 120
Db 22 -----GPLSAELVQKQKDLATATTKDRVTHCTICENTVAQSLRSPF 66
QY 121 QKLGIAEMLFLLCSDDAESDVRWVADECLNKVIRKALMDSNLPRLQLELYKEIKNGAPR 180
Db 67 QKLGIAEMLFLLCSDDAESDVRWVADECLNKVIRKALMDSNLPRLQLELYKEIKNGAPR 126
QY 181 SLRAALWRRFAELAHVLPKCRPYLVNLLPCLTRTSKRPEESVQETLAAAVPKIMASFCN 240
Db 127 SLRAALWRRFAELAHVLPKCRPYLVNLLPCLTRTSKRPEESVQETLAAAVPKIMASFCN 186

QY 241 FANDNEIKVLLKAFIANLKSSPTIRRTAGSNAVSTICQHSRRTOFYXWLLNVLLGLVLP 300
Db 187 FANDGETKMLLKAFVANLKSSPTIRRTAASSAVSVCQHSRRTHFYTWLLNVLLGLVLP 246
QY 301 VEDEHSTLLTGLVLLRLYLVLPLLOOVKDTSLKSGFGVTRKEMEYPSAEOLVQVYELT 360
Db 247 VDEHSHLLTGLVLLRLYLVLPLLOOVKDTSLKSGFGVTRKEMEYPSAEOLVQVYELT 306
QY 361 LHHTQHNDHNVVTALELLQQLFRTPELLOTLITAVGGIGQLTAAKESGGRSRSSTV 420
Db 307 LHHTQHNDHNVVTALELLQQLFRTPELLOTLITAVGGIGQLTAAKESGGRSRSSTV 366
QY 421 ELIAGGSSCPVLRSKOKKVLGGBEALDDSESDVSSSALITASYKDEISGELAAS 480
Db 367 ELIA-GGSTCPLLRKQKGLLSGEEGLDPE-RAEVTTSFTASVGGSSSEAPS 424
QY 481 SGVSTPGSAGHDITEOPR-SQHTLO-ADSVDLASCDL---TSSATDGDDEEDILSHSSQ 535
Db 425 SGVSSLGTS--DIITEQPSRQHALQGDSDVLSASEQGVGPTPDEEEDMLSRSSG 482
QY 536 VSAVPS---DPAMDND---GTOASSPISDSQTTTEGPDSPATVPSDSSIVLDGTDNQY 589
Db 483 GAGLVSTSGDLVTDANQMSAGAVSSSPSSQTTTEGPDSPATVPSDCAELVLDGSESQY 542
QY 590 LGLOIQPQDEDEATGILPDEASEAFRNSMALQOAHLLKNHSHCROPSDSSVDKFEVLR 649
Db 543 SGMQIGTLQDEEEGSAAPPDPKPEPFSQSALALSKPHLLGEGHNNQSSDSSVDKFEVLR 602
QY 650 DEATEPGDQENKPRKIGDQSGDDSDAPLVHCVRLLSASFLLTGKKNVLPDROVRVS 709
Db 603 EEVLEPAELDNKPRKIGDGHYTDKPEELMHCVRLLSASFLLTGKKNVLPDROVRVS 662
QY 710 VKALALSCVGAVALHPESFSLKYLKVPDLTTEPEBQYVSDILNYIDHGDQVQRGATAI 769
Db 663 VKALAVSCVGAVALHPESFSLKYLKVPDLTTEPEBQYVSDILNYIDHGDQVQRGATAI 722
QY 770 LCGTILCSILSRFHVGDMGTTRTGTNTGTFSLADGICPLRLKTKDESSVTKCLACTAV 829
Db 723 LCGALIOAIQKTRNTETWLSQISQVTSQVSSVTLENFVPLLOSLKDESSVTKCMCAAV 782
QY 830 RNCVNLSCSSYSELGLQIIVDITLNRNSVWLVRTELLETABIDPRLVSFLKAEKBNL 889
Db 783 RHCIMALCNGSLSELGLQIIVDITLNRNSVWLVRTELLETABIDPRLVSFLKAEKBNL 842
QY 890 HRGAHYTGILKQRLVNLVNIHLGDDEPRVHVAASLIRLVPLKFKCDQGDQV 949
Db 843 HKGEHYTGILKQRLVNLVNIHLGDDEPRVHVAASLIRLVPLKFKCDQGDQV 902
QY 950 VAVARDQSVVLKLLMHETQPPSHFSVTITRIYRGYNLPSITDVTMNNLSRVIAVS 1009
Db 903 VAIARDQSVVLKLLMHETQPPSHFSVTITRIYRGYNLPSITDVTMNNLSRVIAVS 962
QY 1010 HELITSTRALTFCCEALCLLSTAFPCVINSGLWHGCVPLS----- 1052
Db 963 HALTSSTSRAMTFCCEALCLLSTAFPCVINSGLWHGCVPLS----- 1022
QY 1053 -----ADSESRKSTCVGMATMILTLSSAWFPLDLSAQHALLAGNLLAASAPKSLRS 1106
Db 1023 SLUSQSGSNEEVRSITVGVASVNLSSIAFWFPLDLSAQHALLAGNLLAASAPKSLRS 1082
QY 1107 SWASEEANPAATKOEVPWALGDRALVPMVEQLFSLHLLKVINCAHVLDDVAVGPAIKA 1166
Db 1083 PWAGEESSPASSKVEEPWALNDRLSVVMBQLFSLHLLKVINCAHVLDDVAVGPAIKA 1142
QY 1167 ALPSLTNPSPSLPIRRKKEKEPEGOASVPLSPKKGSEASASROSDTSGPVTTSKSSSL 1226
Db 1143 SLPSLTNPSPSLPIRRKKEKEPEGOASVPLSPKKGSEASASROSDTSGPVTTSKSSSL 1201
QY 1227 GSFYHLPYSLKHLVDLKHATHYKVTLDLQNSTEFGGFLRSALDVLSQLLELATLQD 1286
Db 1202 GSFYHLPYSLKHLVDLKHATHYKVTLDLQNSTEFGGFLRSALDVLSQLLELATLQD 1261
QY 1287 KCVEEILGLYKSCFSREPMMATVVCVQQLKTLFTGNTLASQFDGLSSNFSKSGRAQLRGS 1346

Db 1262 KCVEEILGLYKSCFSREPMMATVVCVQQLKTLFTGNTLASQFDGLSSNFSKSGRAQLRGS 1321
QY 1347 SSVRGLYHYCFMAYPTHTFTQALADASLRNVQAEQENDTSGWDFVQKYSTQKTNLTS 1406
Db 1322 SSVRGLYHYCFMAYPTHTFTQALADASLRNVQAEQENDTSGWDFVQKYSTQKTNLTS 1381
QY 1407 VTKNADKNATHNHLRLEPLVIRKALKYTTTTCVQKQVLDLLAQLVQLRVNVYCLLDS 1466
Db 1382 VTKNADKNATHNHLRLEPLVIRKALKYTTTTCVQKQVLDLLAQLVQLRVNVYCLLDS 1441
QY 1467 DOVFIFGVLFKOFEXIEVQGFRESEAIIPNIFFLVLLSYERYHSHQIITGPKIOLCDGI 1526
Db 1442 DOVFIFGVLFKOFEXIEVQGFRESEAIIPNIFFLVLLSYERYHSHQIITGPKIOLCDGI 1501
QY 1527 MASGRKA 1533
Db 1502 MASGRKA 1508
RESULT 4
QYV3N4
ID QYV3N4 PRELIMINARY: PRT: 3583 AA.
AC QYV3N4;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Huntingtin protein.
GN HUNTINGTIN OR HSAP/HD OR CG9995.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RX [1]
SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.F.,
RA Brandon R.C., Rogers J.H.C., Blazer R.G., Champagne M., Pfeiffer B.D.,
RA Wan K.H., Doyle E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Frankoch C., Baldwin D.,
RA Ballwe R.M., Basu A., Baxendale J., Bayraktaroglu L., Bessley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brodeur P., Brothier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlike C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostler C., Gong F., Gabor N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Correll J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Markulov G., Milshina N.V., Moberly C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Pittman K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Sier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirski R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,


```

QY 1335 SKSQRAQRLGSSVRPGLYHYCFMAY-----THFTQALADASLRNM 1377
D 1757 SATGGNGSEIG-----HHAFAFRPYFAAKGRHGASSTLLPTINSKPAVAVGSORGA 1808
QY 1378 -VQAEQNDTSGWFDVLQKVSTQLKNTLVTKNRADKNAIHNHRLFPFLVIAKALKQYT 1436
D 1809 PTDARQPIDAG-----PLQDMGLMFVHGLQPPPTPAGDCVRL---IKLFEPYVIYCLTFLM 1861
QY 1437 TTTCTVOLQKQVLDLLAQLVQLRVNVCYLLSDQVFGFVLKQFEYIEVGFRESEALIIPNI 1496
D 1862 KSNAL-VQAPILRLSLQLLDLVNTYSILDSKNVIFQVLSNMDLIEGGIDRANAFIMVPPM 1920
QY 1497 FFEVLVLSYRYSKOIIGIPKTIQCLDGIMASG 1530
D 1921 LRFELVQLTHK--SDROLIPIKIISITNLLANG 1952

RESULT 5
Q90999 PRELIMINARY; PRT; 3584 AA.
AC Q90999;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE HUNTINGTIN homolog.
GN HUNTINGTIN OR CG9995.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OC NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Takano H., Bernards A., Gusella J.F.;
RT "Drosophila HD homolog.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AFI77386; AAD51369.1; -.
DR FlyBase; FBgn0027655; huntingtin.
DR InterPro; IPR000091; Huntingtin.
DR InterPro; IPR000834; Zn_Carboxypept.
DR Pfam; PF03541; Huntingtin; 1.
DR PROSITE; PS00133; CARBOXYPEPT_ZN_2; UNKNOWN_1.
SQ SEQUENCE 3584 AA; 395854 MW; DB7C7FB764BCA6E CRC64;

Query Match 8.5%; Score 668; DB 5; Length 3584;
Best Local Similarity 19.6%; Pred. No. 9.3e-32;
Matches 406; Conservative 254; Mismatches 642; Indels 772; Gaps 62;

QY 89 KELSATKRDVNHCLTICENIVAQSVRNSPEFQKLLGIAMELFLLCSDDAESDVRVMADE 148
D 20 RNTFCSQKQITCFQQAECIMSPSLAGHINYAAHCGTATNVLVLLFCEDVDVSVRMSAE 79
QY 149 CLNKVIALMDSNLPRLQLELYKEIKKNGAPRSRLAALWRFELAHVLRPKRPRYLVLN 208
D 80 NLNKLRSLEKTRVSRITMDLGEIKRNGNQRSLCNLFYSYAPQIKRKHKKWYAVRL 139
QY 209 LPLCTRTRSKRPEESVQETLAAAVPKIMASFGNFANDEIKVLLKAFIANLKSSSPTRT 268
D 140 LQCMTTTSQRKTLQQLTQCFVHFRSHRTOQLGSDESCKLFETFLDQISSCAVRRRC 199
QY 269 AAGSAVSIQHSRRTOFYVSWLLNVLGLLVPVEDEHSTLLILGLVLTLLRYLPLPQQOV 328
D 200 SAQCMSLIENARNRSLMARHGVNRYVMEELLTLDQANS---VLGALLRLRLPQLIRGY 256
QY 329 KDTSLKSGFVTRKEME-----VPSABQLVQVYELFLH--HTQHDHNVVYTGAELELQOL 382
D 257 PGDSHEDSESISLAKKQKQKQKQTTTSDCRQIIEYDYCLHLLSTQHTANHAJINA----- 309
QY 383 FRTPPPELLOTLTAVGGIGO-LTAAKESGGRSGRSIGIVELIAGGSSCSPLSRKO--- 438
D 310 -----TLEVINGILQAVDAASDQCQSLSLQSLRQLLNCNOQLHNEYLRKRSLK 359

```

```

QY 439 -----KGKVLGBEEALEDDSESR--SDYSSSAULTASVKDEISGELA----- 478
D 360 NOIFOLKNYEVATSOHOLEDEDEDEDEVDVVGATAMQMKKNSNAKLOQAKCRQOQHOH 419
QY 479 -----ASSGVSTPGSAGHDIITEQPRS----- 500
D 420 QOQLEVDNSSLGI---NAGEDAPTEAPSSVADEGGPESTKLCHIRNAARSISECVAS 475
QY 501 -----QHTLQADSVDLASCOLTSSATDCDEED--ILSHSSSVQSVAPSDPMDLND 549
D 476 DEDKQGGHROORDEGCVVVAEDDDDDDDDDDDMELLSAECDDFTL-----SQLNE 530
QY 550 GTQASSPISDSQTTTEGDSAVTPSDSSEIVLD-----GTD--N 587
D 531 QOQALSAALKLPTTTAASSGGAATSQDDKLIDVDADVGGLPKQHQSSLQNLLAGSDOKS 590
QY 588 QYLG-----LQIGQPOD---EDEEATGI----- 607
D 591 QHLSIDIDNESFNSIDFAEITAGSKQEQOQHPADDSVESGDATAIGTFPNNLLSHNA 650
QY 608 LPDEASEAFRNS-----SMALQQAHLKKNMSHC-----RQP- 638
D 651 ASESVSKLFROSSGSKSTPKSASTPAPADKSDAISASLTLSLTSASSNLEPPERQPL 710
QY 639 -----SDS-----SVDKFVLRDEATEPGD-----ENKPCR 664
D 711 TAETPTPEVDECSITASTASTALMDAPAVEAASKPETPQLRGTPNANPFLVENSPLR 770
QY 665 -----IKDGIGOSTDDSNAPLVHCVRLLSASFLLTGKKNVLVDPDRDVSVKALALS 716
D 771 QTVVGRALITVKIGSIL--EOSLVYITARLVAARFLSSGQAAGLQDPSISRVSKISLSLA 828
QY 717 CVGAVALHP-----ESFFSKL-----YKVP 738
D 829 VTAQCVRLAPKTLQSLSEISQELQLLEATSQICSGDSTQVSSPQSDNSQVGEXPL 888
QY 739 DTETVP----- 744
D 889 DSSLVPTSLLENLLLDIKDDHFGPSTCPAYLQSATPTLSRSDASAVLLEVTSSRSA 948
QY 745 ----- 744
D 949 KKEEMLSKSEIIESSRYPTVAVEDVPPSPMPRPKTKSTRSRVGLGTSSITSSSP 1008
QY 745 -BEQYVSDILNYIDHGPQVRGATAILCGLTICS----- 777
D 1009 QSROKLSDLLFHCHDCLPILRGVQVQVGNFLQSSGAGLFLDLQRLGLQHLAAILKGF 1068
QY 778 -----ILSR-----SRFHVGDW----- 790
D 1069 EDEIHTVVIQALNADFDFKIPNVVSKYLTEPPCHYHAHQOQOQOQOQOQOQOQOQOQOQ 1128
QY 791 -----GTIRLTGNTF-----SLADCIPLLR 811
D 1129 LQHSQGGQKRSQGAQTGQQTFAKQDQNALSSQOQRRPNDACTCANSATDNDELLA 1188
QY 812 KTLKDESSVTCKLACTAVRN-----CVMSLCSSSYSELGLQ 847
D 1189 ALLND-----FQLOSTGMRQOQKNNSTDIGOSGNEPDLPPNPAVAEPFCVFAISP---K 1240
QY 848 LIIDVLT-LRNSSYMLVTRTELLETAEIDFRLVSFLEAK-----AENLHRGAHHYTG----- 898
D 1241 LLLSKRLCHHKNKYLQNKYAEVISNLNYLLRSYANFRCAIDNKNKGARKQSDKWWP 1300
QY 899 -----LLKQERVNLNVVTHLQDEDEPRVRHVAASLIRLV----- 934
D 1301 MDASSVCHSVRDAEGEDIVCTYEAQFLAELLHLLGDDARVREHAACCLCFIMQTARQD 1360
QY 935 PKLFYKCOGOADPV-----VAVARDQS-----SVYKLKLMHETQ--- 969
D 1361 PSODQAGGGGGGDIENGNVNVETQOTNFNLLWDFDYRIFRFGSMVTLRNLFRASSIV 1420
QY 970 PP-----SHFVSTITIRYRG 985

```

```

Db 1421 PLALDALATNSAPSYDPTGSGSTSTSSAGSGSAAVSAASAYFEASYGIGIAEG 1480
QY 986 YNL-LPSITD---VTMNNLSVIAAASHLITSTRALTTCGCCALCLITSTAPPVCIWS 1041
Db 1481 HVPALASASORQIAQEKVLAKVLRNKLMTLNDKNVQFGIIVYALRLLRHFNFDYQ 1540
QY 1042 LGWHGCVPLSASDESRSKCTVGMATMLTLTSSAWF-----PLDLSAHQDALILAGNLLA 1097
Db 1541 QAWL-----EENFEICISYAYINNATAADLGCCNDLIDVNGKMA 1581
QY 1098 ASAPKSLRSSWASEANPAATKQEVVPALGDRLVPMVQOLFSLHLKLVINICAHVLD 1157
Db 1582 -----GAMUSSGEPNTA-----HLDFLLRHVSVMKMLNIYHVLTN 1615
QY 1158 VAPGAIPAALPLTNPPSLSPIRKKGKEPEQASVPLSPKKGSEASASROSQDTS 1217
Db 1616 QRPPTA-----GSGSSSKQPKSELFAREQPAAT--- 1646
QY 1218 VTTSSSSLGSPYHPLSVLKLHVLKATHANYKVTLDLQNSTEKFGFLRSALDVL 1277
Db 1647 -----LQALGVFAGDVVYMKLYNLRGANDSYKITTIN-QEAGSLICLLKTLHA 1700
QY 1278 E---LATLQDICKVEEILGYLKSCFSREPMMATVVCVQQLKTLFTGTNLASQF 1334
Db 1701 EGMASAPPEL-KLIEEILHLVLRILNYAPAECAVCLRLKYLFAQNYASQ---VR 1756
QY 1335 SKSOGRAORLGSVVRPGLYHYCFMAY-----THFTQALADASLRNM 1377
Db 1757 SAG-----IGNGTGIG-HIAAFMRPFAAKGRGHGASSLTLLPTINSKPAVAV 1809
QY 1378 -VOAEQNDTSQFVQVLOKVTQTLNLTSTKNRADKNAIHNIRLPEPLVIRKALK 1436
Db 1810 PTDARQPIDAG---PLQDMGMFLVHGLQPPTPAGDCVRL---IKLPEPVIYCL 1862
QY 1437 TTTCVOLQKQVLDLAQLVOLRVNVCCLDSQVFGFVLFKQFIEVQCFRESEAI 1496
Db 1863 KSNAL-VQAPILRLSLQDLNVTYILDSKNVIFDQVLSNMDLIEGGIDRNAP 1921
QY 1497 FFFVLVLSYRHSKQIIGIKPIIQLCDGIMASG 1530
Db 1922 LRFLVOLTHK--SDRLITIPKISITNLLANG 1953

RESULT 6
O18897
ID O18897 PRELIMINARY; PRT; 75 AA.
AC O18897;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DE 01-JAN-1998 (Tremblrel. 05, Last annotation update)
GN HD.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RA Shiba H., O'Brien D.P., Chen Y.-W., Johnson G.S.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF029754; A0884207.1;
FT NON_TER 1
FT NON_TER 75
SQ SEQUENCE 75 AA; 8133 MW; ADD779011B8CD367 CRC64;

Query Match 4.5%; Score 358; DB 6; Length 75;
Best Local Similarity 94.7%; Pred. No. 3,9e-15;
Matches 71; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

QY 670 GQSTDDSDAPLVHCVRLLSASFLLTGKKNVLPDRDVRVSVKALALSCVGAVALHPESF 729
Db 1 GQSTDDSDVPLVHCVRLLCASFLLTGKKNVLPDRDVRVSVKALALSCVGAVALHPESF 60

```

```

QY 730 FSKLYKVPDLDTTEYP 744
Db 61 FSKLYKVPDLDTMEYP 75

```

RESULT 7

```

Q8T069
ID Q8T069 PRELIMINARY; PRT; 620 AA.
AC Q8T069;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE LD23533P.
GN HUNTINGTIN.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY069515; AAL39660.1;
SQ SEQUENCE 620 AA; 69206 MW; 87F2240FD15C2432 CRC64;

```

```

Query Match 4.5%; Score 356.5; DB 5; Length 620;
Best Local Similarity 22.5%; Pred. No. 1.2e-13;
Matches 137; Conservative 116; Mismatches 268; Indels 89; Gaps 14;

```

```

QY 89 KELSATKKRVNCHLTICENIVASVRNPFQKLLGIAMLEFLCSDDAESDVRVMADE 148
Db 20 RNTECSQKQKTCFOQTAECIMSPSLAGHINYAAHCGTATNVLLFCEDVDVSRMSAE 79
QY 149 CLNKVITKALMDSNLPRLQLELYKEIKNGKNGAPRSALRALMRFABLAHLRPKQCR 208
Db 80 NLAKILRSLEKTRVSRILMDYGEIKRNGNQRSLICNLFSYVAPQIKHKKYAVRL 139
QY 209 LPLCTRTSKRPEESVQETLAAAVPKIMASPGNFANDNEIKVLKAFIANLKSSPTIR 268
Db 140 LQMTTISQRKLTLLQETLQDFYKHFSRHTQQGLSDSECKLFFETLQIISDCAVRR 199
QY 269 AAGSAVSICQHSRRQTQVYFVSWLLNVLGLLPVVEDEHSTLLILGLVLLTLYLVPL 328
Db 200 SAQNCMSLIENARNRSLMARHGVNKVYVWELLTQOQANS---VLGALGLRLLLP 256
QY 329 KDTSLKSGFVTRKEME---VSPSAELVQVYELTLH---HTQHDHNVVVTGALEL 382
Db 257 PGDSHEDSESLAGKQKQOQOQTTSDCRQIIEYDYLHLLSTQHTANHAINA----- 309
QY 383 FRTPPELQLTAVGGIGQ-LTAAKEESGSRSSRSISVELIAGCGSSSPVLSRKQ--- 438
Db 310 -----TLEVINGLQAVDAASDGGQCSQSLGRQLCNOQLQHNEYLRRLRSK 359
QY 439 -----KGKVLGEEALEDDSESRSDDVSSSALTASVKDEISGELAASSGVSTP 492
Db 360 NQIFQLKNYEVATSQHOLEDENE-----DYDELVVGATAMQMKNSNAKLOQA 409
QY 493 IITEQPRSQHTLQADSVDLASCDLTSSATDGEDDEILSHSSSSQSVAVSPDAMD 552
Db 410 KCREQOQHQQOOLE-VDNSLGINAG-----EDAPTEAPSSVADEGEP----- 457
QY 553 ASSPISDSSQTTTE-----GPDSSAVTPDSSSIIVLDTGDNQVILGICIQOPDE 603
Db 458 LRCHIRNAASICEVASEDEDKQGGHQROREDGYYVAEDDDDD-----DDDDDD 510
QY 604 ATGILPDEASEAPRNSMALQQAHLKKNHSHCHQPSDSSVDKFLVRDEATEPG--- 660

```



```
Db      30 -----PPPPPPPLPGP-AEELHPR 67

RESULT 10
Q91TWI PRELIMINARY; PRT; 729 AA.

ID Q91TWI AC Q91TWI;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Similar to HCMVA US23 and MCMVS M43.
OS Tupaia herpesvirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae.
OX NCBI_TaxID=10397;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2;
RX MEDLINE=21211637; PubMed=11312357;
RA Bahr U., Darai G.;
RT "Analysis and Characterization of the Complete Genome of Tupaia (Tree Shrew) Herpesvirus."
RL J. Virol. 75:4854-4870(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=2;
RA Darai G., Bahr U.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
DR ENBL; AF281817; AAK57026.1; -.
DR InterPro; IPR003360; US22.
DR Pfam; PF02993; US22; 1.
SQ SEQUENCE 729 AA; 81192 MW; D5FDE57AABFB3EAB5 CRC64;

Query Match          2.9%; Score 229.5; DB 12; Length 729;
Best Local Similarity 48.5%; Pred. No. le-05;
Matches 48; Conservative 6; Mismatches 24; Indels 21; Gaps

QY 18 QQQQQQQQQQQQQQQQQQQQQQQQPPPP-----PPPPPPPOLQPQP-----QA 57
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 509 QQQQQQQQQQQQQQQQQQQQQQQQQQQPPPPKQKQPPRPPPSRFSEERPEEPSEDSPILSS 568

QY 58 QPLLQPQPPPPPPPPPPGPAVAAEPLHRPKKLSATKK 96
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 569 SPIQPVPVTIPPPPPPPP-PAFHDHDCASPSELAVLER 606

RESULT 11
Q9BQQ4 PRELIMINARY; PRT; 1214 AA.

ID Q9BQQ4 AC Q9BQQ4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE KIAA0298 protein.
GN KIAA0298
GS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Amid C., Hankeln T., Winterpacht A., Zabel B., Schmidt E.;
RT "Comparative sequencing of the human chromosome 11p15 and mouse chromosome 7."
RL Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
DR ENBL; AJ400879; CAC35389.1; -.
DR HSSP; Q92831; 1B91.
DR InterPro; IPR003649; Bbox_C.
DR InterPro; IPR001487; Bromodomain.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR InterPro; IPR002965; P_rich_extensn.
DR InterPro; IPR000315; ZnF_Box.
DR InterPro; IPR001965; ZnF_PHD.
```


[illegible]

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 13:58:08 ; Search time 85,9798 Seconds
(without alignments)
2391.326 Million cell updates/sec

Title: US-09-904-987-7
Perfect score: 7892
Sequence: 1 MATLEKLMKAFESLKSFOQQ.....DGIMASGRKSPQYRLCSP 1543

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7821.5	99.1	3144	18 AAW09871	Human huntingtin.
2	7821.5	99.1	3144	20 AAY33493	Human huntingtin p
3	7821.5	99.1	3223	22 ABB11407	Human huntingtin's
4	7821.5	99.1	3223	22 ABB11470	Human huntingtin's
5	7805.5	98.9	3144	15 AAR58777	Protein encoded by
6	7805.5	98.9	3144	18 AAW36887	Previously undescri
7	7805.5	98.9	3144	19 AAW44742	Human huntingtin p
8	7145.5	90.5	3119	18 AAW36888	Mouse huntingtin s
9	7145.5	90.5	3119	18 AAW44743	Mouse huntingtin p
10	7041	89.2	3139	20 AAY08898	Human huntingtin's

11	2967.5	37.6	589	20 AAY33503	Human apopain clea
12	2784.5	35.3	552	20 AAY33502	Human apopain clea
13	2676.5	33.9	530	20 AAY33501	Human apopain clea
14	2589.5	32.8	513	20 AAY33500	Human huntingtin p
15	1974	25.0	418	18 AAW26545	Mouse huntingtin's
16	884.5	11.2	171	20 AAW99022	Human huntingtin p
17	674.5	8.5	3583	22 ABB4814	Drosophila melanog
18	456	5.8	108	20 AAW95071	Amino acid sequenc
19	456	5.8	108	20 AAW95071	Amino acid sequenc
20	416.5	5.3	79	22 AAB69616	Huntingtin's disea
21	213.5	2.7	86	20 AAW95073	GST-HD fusion prot
22	213.5	2.7	86	20 AAW95073	GST-HD fusion prot
23	210	2.7	94	20 AAW95075	GST-HD fusion prot
24	210	2.7	94	20 AAW95075	GST-HD fusion prot
25	210	2.7	192	22 AAB48292	GST-HD fusion prot
26	209	2.6	55	20 AAW95072	Human zF4 protein.
27	209	2.6	55	20 AAW95072	GST-HD fusion prot
28	208	2.6	63	20 AAW95074	GST-HD fusion prot
29	208	2.6	63	20 AAW95074	GST-HD fusion prot
30	208	2.6	197	21 AAY83079	F-box protein FBP-
31	205	2.6	3536	22 ABB65480	Drosophila melanog
32	201.5	2.5	1954	22 ABB59197	Drosophila melanog
33	195.5	2.5	2783	13 AAR23962	AFP-1. Homo sapie
34	195.5	2.5	2783	13 AAR23963	AFP-1 (Ala 2480 Va
35	195.5	2.5	2783	22 AAB2946	Human androgen rec
36	194	2.5	99	22 AAO02076	Human polyptide
37	193	2.4	85	22 AAO04573	Human polyptide
38	193	2.4	995	22 ABB62510	Drosophila melanog
39	190	2.4	80	22 AAO04412	Human polyptide
40	188	2.4	1368	22 AAB60262	Drosophila melanog
41	186.5	2.4	98	22 AAB69610	Huntingtin accumu
42	185.5	2.4	1151	22 ABB61598	Drosophila melanog
43	185	2.3	89	22 AAB69608	Huntingtin accumu
44	185	2.3	121	22 AAB69609	Huntingtin accumu
45	185	2.3	123	22 AAB69611	Huntingtin accumu

ALIGNMENTS

RESULT 1
AAW09871
ID AAW09871 standard; Protein; 3144 AA.
XX
AC AAW09871;
XX
DT 27-JUL-1997 (first entry)
XX
DE Human huntingtin.
XX
KW Huntingtin associated protein-1; HAP1; Huntington's disease.
XX
OS Homo sapiens.
XX
FH Key
FT Protein
FT 1..230
FT /note= "Claim 20"
XX
PN WO9717443-A1.
XX
PD 15-MAY-1997.
XX
PF 08-NOV-1996; 96WO-US17858.
XX
PR 09-NOV-1995; 95US-0556419.
XX
PA (UYJO) UNIV JOHNS HOPKINS.
XX
PI Lanahan A, Li S, Li X, Ross CA, Sharp AH, Snyder S;
PI Worley PF;
XX
DR WPI; 1997-281032/25.
XX

PT Determination of the binding of huntingtin to huntingtin-associated
 PT protein-1 - useful for screening for drugs for treating or
 PT preventing Huntington's disease
 XX

Claim 20; Page 22-32; 69pp; English.

XX Human huntingtin (Hn) polypeptide (AAW09871) is the product of the
 CC Huntington's disease (HD) locus. Proteins which specifically bind
 CC to Hn, such as human huntingtin associated protein-1 (HAP1) (see
 CC also AAW09870), can be used in assays for screening drug candidates.
 CC The binding between Hn and HAP1 is enhanced by an expanded
 CC polyglutamine repeat in Hn, the length of which correlates with the
 CC time of disease onset. HAP1, in contrast to Hn, is expressed
 CC selectively in the brain, suggesting that it may contribute to the
 CC brain-specific pathology of HD. Hn, or portions of it, esp. amino
 CC acids 1-230, or yeast cells expressing Hn, can be used to identify
 CC cpds. that bind to, displace or prevent binding of Hn and HAP1.
 CC The method is useful for screening candidate drugs for treating,
 CC delaying onset of, or preventing HD.
 XX

SQ Sequence 3144 AA;

Query Match 99.1%; Score 7821.5; DB 18; Length 3144;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 1533; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 1 MATLEKLMKAFESLKSF---QQQQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPQA 57
 DB 1 MATLEKLMKAFESLKSFQQQQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPQA 60
 QY 58 QPLLPPQPP 117
 DB 61 QPLLPPQPP 120
 QY 118 PEQKLLGTAMEFLFICSDDAEDVRWADECLNKVKKALMDSNLPRLQLELYKEIKNG 177
 DB 121 PEQKLLGTAMEFLFICSDDAEDVRWADECLNKVKKALMDSNLPRLQLELYKEIKNG 180
 QY 178 APRSLAALWRFALWFAHLVLPKCRYPVNLPLCLTRTSKRPEESVQETLAAAPVKIMAS 237
 DB 181 APRSLAALWRFALWFAHLVLPKCRYPVNLPLCLTRTSKRPEESVQETLAAAPVKIMAS 240
 QY 238 FGNFANDNEIKVLLKAFIANLKSSSTIRRTAAGSAVSTCOHSRRTPQYFYSWLLNVLLGL 297
 DB 241 FGNFANDNEIKVLLKAFIANLKSSSTIRRTAAGSAVSTCOHSRRTPQYFYSWLLNVLLGL 300
 QY 298 LVPVEDEHSTLLILGLVLLTLRYLVLPLQQQVKTSLKGSFGVTRKEMEYSPSAEQLVQVY 357
 DB 301 LVPVEDEHSTLLILGLVLLTLRYLVLPLQQQVKTSLKGSFGVTRKEMEYSPSAEQLVQVY 360
 QY 358 ELTLHTHTQHDHNVVGTGALELLOQLFRTPPPELLQTLTAVGGIGQLTAKEESGGRSRSG 417
 DB 361 ELTLHTHTQHDHNVVGTGALELLOQLFRTPPPELLQTLTAVGGIGQLTAKEESGGRSRSG 420
 QY 418 SIVELIAGGSSCPVLSRKQKGVLLGEEALEDDSDRSVSSSALTATSVKDEISGEL 477
 DB 421 SIVELIAGGSSCPVLSRKQKGVLLGEEALEDDSDRSVSSSALTATSVKDEISGEL 480
 QY 478 AASSGVSTPGSAGHDITEQPRSHQTLQADSDVLDASCDLTSSATGDDEEDILSHSSQVS 537
 DB 481 AASSGVSTPGSAGHDITEQPRSHQTLQADSDVLDASCDLTSSATGDDEEDILSHSSQVS 540
 QY 538 AVPSDPAAMDNDGTQASSPISDSSQTTEGPDSAVTPSDSSEIVLDGTDNQVLQIGQP 597
 DB 541 AVPSDPAAMDNDGTQASSPISDSSQTTEGPDSAVTPSDSSEIVLDGTDNQVLQIGQP 600
 QY 598 QDEDEEATGILPDEASEAFRNSSMALQQAHLKNMCHCRQPSDSSVYDKFVLDEATEPGD 657
 DB 601 QDEDEEATGILPDEASEAFRNSSMALQQAHLKNMCHCRQPSDSSVYDKFVLDEATEPGD 660
 QY 658 QENKPCRIKIDIGQSTDDSDAPLVHCVRLLSASFLLTGTGKNVLPDPDRVRSVKALALSC 717
 DB 661 QENKPCRIKIDIGQSTDDSDAPLVHCVRLLSASFLLTGTGKNVLPDPDRVRSVKALALSC 720

QY 718 VGAVALHPSEFFSKLYKVPDLTTEYPPEQVSDILNYIDHGPQVRGATAILCGTLICS 777
 DB 721 VGAVALHPSEFFSKLYKVPDLTTEYPPEQVSDILNYIDHGPQVRGATAILCGTLICS 780
 QY 778 ILSRSRHFVGDWMTGIRTLTGNTFSLADCIPLLRKTKLKDESSVTKLACTAVRNCVMSLC 837
 DB 781 ILSRSRHFVGDWMTGIRTLTGNTFSLADCIPLLRKTKLKDESSVTKLACTAVRNCVMSLC 840
 QY 838 SSSSELGLOLIIDVLTNRSSYWLVRTELETLAEIDFRLVSELEAKAENLRGAHHYT 897
 DB 841 SSSSELGLOLIIDVLTNRSSYWLVRTELETLAEIDFRLVSELEAKAENLRGAHHYT 900
 QY 898 GLLKLQERVLNNVVIHLLGDDEPRVRHVAASLRLVKLFYKDCQDQADPVVAVARDQS 957
 DB 901 GLLKLQERVLNNVVIHLLGDDEPRVRHVAASLRLVKLFYKDCQDQADPVVAVARDQS 960
 QY 958 SVYLKLLMHETQPPSHFSVSTITRIYRGYNLLPSITDVTMNNLSRVIAVSHELITSTT 1017
 DB 961 SVYLKLLMHETQPPSHFSVSTITRIYRGYNLLPSITDVTMNNLSRVIAVSHELITSTT 1020
 QY 1018 RALTFGCEALCLLSTAFPVCISLWGHGVPPLSASDESRSKCTVGMATMILTLLSSAW 1077
 DB 1021 RALTFGCEALCLLSTAFPVCISLWGHGVPPLSASDESRSKCTVGMATMILTLLSSAW 1080
 QY 1078 FPLDLSAHQDALILAGNLLAASAPKSLRSSWASSEANPAATKQEEVMPALGDRLVPMV 1137
 DB 1081 FPLDLSAHQDALILAGNLLAASAPKSLRSSWASSEANPAATKQEEVMPALGDRLVPMV 1140
 QY 1138 EQLFSHLLKVINICAHVLDVAVPGPAIKALPSLTNPSPSLPIRRKKEKEPEGQASVPL 1197
 DB 1141 EQLFSHLLKVINICAHVLDVAVPGPAIKALPSLTNPSPSLPIRRKKEKEPEGQASVPL 1200
 QY 1198 SPKKGSEASASRSQSDTSGPVTTSKSSLSGFSYHLPYSLKLDVLCATHANYKVTLDQN 1257
 DB 1201 SPKKGSEASASRSQSDTSGPVTTSKSSLSGFSYHLPYSLKLDVLCATHANYKVTLDQN 1260
 QY 1258 STEKFGGFLRSALDVLISOILELATLQDQKVEEILGYLSCFSRPMMATVCVQQLLKT 1317
 DB 1261 STEKFGGFLRSALDVLISOILELATLQDQKVEEILGYLSCFSRPMMATVCVQQLLKT 1320
 QY 1318 LFGTNLASOFGGLSSNPSKSGRAQLGSSSRVPLGYHYCFMAYPTHFTOALADASLRNM 1377
 DB 1321 LFGTNLASOFGGLSSNPSKSGRAQLGSSSRVPLGYHYCFMAYPTHFTOALADASLRNM 1380
 QY 1378 VQAEQENDTSGDFVLQKVSTQLKTNLTSTVTKNRADKNAIHNHRLFEPLVVKALKQYTT 1437
 DB 1381 VQAEQENDTSGDFVLQKVSTQLKTNLTSTVTKNRADKNAIHNHRLFEPLVVKALKQYTT 1440
 QY 1438 TTCVOLQKQVLDLLAQVLQVRNYNCLLSDQVFTGFVLKQFEYTEVQGFRESEAIIPNIF 1497
 DB 1441 TTCVOLQKQVLDLLAQVLQVRNYNCLLSDQVFTGFVLKQFEYTEVQGFRESEAIIPNIF 1500
 QY 1498 FFLVLLSYERYHSHKQIIGIPKIIQLCDGIMASGRKA 1533
 DB 1501 FFLVLLSYERYHSHKQIIGIPKIIQLCDGIMASGRKA 1536

RESULT 2

AAV33493

ID AAV33493 standard; Protein; 3144 AA.

XX

AC AAV33493;

XX

DT 19-JAN-2000 (first entry)

XX

DE Human huntingtin polypeptide.

XX

KW Proapoptotic; dependence domain; p75NTR; androgen receptor; DCC;
 KW huntingtin polypeptide; Machado-Joseph disease; SCAL; SC2; SC6;
 KW atrophin-1; cell death; apoptosis; Huntington's disease; head trauma;
 KW Alzheimer's disease; Kennedy's disease; spinocerebellar ataxia; stroke;
 KW dentatorubropallidolysian atrophy; cell proliferation; cell survival;

QY 1498 FFLVLLSYERYHSKQITGIPKIIQLCDGIMASGRKA 1533
 |||
 |||
 Db 1501 FFLVLLSYERYHSKQITGIPKIIQLCDGIMASGRKA 1536

RESULT 3

ABB11407
ID ABB11407 standard; peptide: 3223 AA.

11-JAN-2002 (first entry)

XX
DE Human Huntington's disease protein homologue. SEQ ID NO:1777.

Human; cytokine; cell proliferation; cell differentiation; growth factor;
 haematopoiesis regulation; tissue growth; immunomodulator; activin;
 inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
 proliferation; metastasis; cancer; tumour; haematopoietic disorder;
 myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
 chronic inflammatory condition; proliferative retinopathy;
 atherosclerosis; coronary heart disease; arterial ischaemia;
 bone disorder; osteoporosis; vascular growth disorder;
 tissue regeneration; wound healing; infection; immune disorder;
 cell culture; drug screening; gene therapy; antiinflammatory;
 antiasthmatic; antiarthritis; haemostatic; antiarteriosclerotic;
 cytostatic; osteopathic; vasotrophic; cardiant; virucide; antibacterial;
 antifungal; vulnerary; antiulcer.

OS Homo sapiens.

XX PN W0200157188-A2

XX
09-AUG-2001XX
PF
05-FEB-2001: 2001WO-US03800.

03-FEB-2000: 2000US-0496914.

PR 27-APR-2000; 2000US-05608/5.
XX

PA (HYSE-) HYSEQ INC.
XX

PI Tang YT, Liu C, Drmanac RT;
XX

DR WPI; 2001-457740/49.
DR N-PSDB: ABA08651

XX
DT
Human protocols and

treating or ameliorating a medical condition in a mammalian subject

XX
PC 101-103. 106

XX

CC sequences ABA08225-ABA09574 represent nu

CC nucleotide of the invention, methods of producing the novel polypeptides
CC antibodies against the polypeptides, methods of detecting the nucleotides
CC or polypeptides in a sample, and methods of identifying compounds which
CC bind to polypeptides of the invention. Although novel, many of the
CC polypeptides of the invention have homology to known proteins, thereby
CC giving an insight into their probable biological activities, and hence
CC potential therapeutic applications. The polypeptides of the invention may
CC have various activities, including cytokine, cell proliferation or cell
CC differentiation activities; stem cell growth factor activity;
CC haematopoiesis regulatory activity; tissue growth activity;
CC immunomodulatory activity; activin- or inhibin-related activities;
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC thrombolytic activities; receptor or ligand activities; or may be
CC involved in oncogenesis, cancer cell proliferation or metastasis.
CC Depending on their biological activities, polypeptides and nucleotides of
CC the invention are useful for preventing, treating or ameliorating medical
CC conditions, e.g., by protein or gene therapy. Such conditions include

Db 800 VGAVALHPSEFSKLYKVPLODTTEPEQYVSDILNYIDHGDPQVRGATAILCGTLLCS 859
QY 778 ILSRSFHVGDWNGTIRTLTGNTESLADCIPIRLKRLTKLXDESSVTCKLACTAVRNCVMSLC 837
Db 860 ILSRSFHVGDWNGTIRTLTGNTESLADCIPIRLKRLTKLXDESSVTCKLACTAVRNCVMSLC 919
QY 838 SSSYSGLGQLIIDLTLRNSYWLVRTELLETLAEIDFRLVSFLEAKAENLHRCAGHYT 897
Db 920 SSSYSGLGQLIIDLTLRNSYWLVRTELLETLAEIDFRLVSFLEAKAENLHRCAGHYT 979
QY 898 GLIKQERVLNVVTHLLGDEDPVRHVAAASLIRLVPKLFYKCDQGGQADPVAVARQDS 957
Db 980 GLIKQERVLNVVTHLLGDEDPVRHVAAASLIRLVPKLFYKCDQGGQADPVAVARQDS 1039
QY 958 SYVLKLLMHETOPPSHFVSTIRYRGYNLLPSITDVTMNNLSRVAAVSHELITTT 1017
Db 1040 SYVLKLLMHETOPPSHFVSTIRYRGYNLLPSITDVTMNNLSRVAAVSHELITTT 1099
QY 1018 RALTEGCCCEALCLLSTAFPVCISLWGHGCVPLPSASDESRSKCTVGMATMLTLTLLSSAW 1077
Db 1100 RALTEGCCCEALCLLSTAFPVCISLWGHGCVPLPSASDESRSKCTVGMATMLTLTLLSSAW 1159
QY 1078 FPLDLSAHQDALILAGNLLAASAPKSLRSSWASEEANAATKQBEVWPALGDRALVPMV 1137
Db 1160 FPLDLSAHQDALILAGNLLAASAPKSLRSSWASEEANAATKQBEVWPALGDRALVPMV 1219
QY 1138 EOLFSLHLKVINICAHVLDVDPAGPAIKAAPLSLTNPPSLSPIRKKGKEPGEQASVPL 1197
Db 1220 EOLFSLHLKVINICAHVLDVDPAGPAIKAAPLSLTNPPSLSPIRKKGKEPGEQASVPL 1279
QY 1198 SPKKGSEASASRSQSDTSGPVTTSKSSLSGFSYHLPSYKLHDLVKATHANKVTTLDLQN 1257
Db 1280 SPKKGSEASASRSQSDTSGPVTTSKSSLSGFSYHLPSYKLHDLVKATHANKVTTLDLQN 1339
QY 1258 STEKEGGFLRSALDVLSQLLELATQDICKVEEILGYLKSFCFSREPMATVVCVQQLLKT 1317
Db 1340 STEKEGGFLRSALDVLSQLLELATQDICKVEEILGYLKSFCFSREPMATVVCVQQLLKT 1399
QY 1318 LFGTNLASQDGLSNPKSQGRAGLSSSVRPLGYHYCFMAYPTHFTQALADASLRNM 1377
Db 1400 LFGTNLASQDGLSNPKSQGRAGLSSSVRPLGYHYCFMAYPTHFTQALADASLRNM 1459
QY 1378 VQAEQENDTSGFVQLQVSTQKLTNLTSVTKNRADKNAIHNLIRLFEPLVVKALKQYTT 1437
Db 1460 VQAEQENDTSGFVQLQVSTQKLTNLTSVTKNRADKNAIHNLIRLFEPLVVKALKQYTT 1519
QY 1438 TTCVOLQKQVLDLLAQLVQLRVNCLLSDQVFFGFLVKQFEYIEVGQFSESEAIIPNIF 1497
Db 1520 TTCVOLQKQVLDLLAQLVQLRVNCLLSDQVFFGFLVKQFEYIEVGQFSESEAIIPNIF 1579
QY 1498 FFLVLLSYERVHSHKQIIGPKIIOQLCDGIMASGRKA 1533
Db 1580 FFLVLLSYERVHSHKQIIGPKIIOQLCDGIMASGRKA 1615
RESULT 4
ABBI1470
ID ABBI1470 standard; peptide; 3223 AA.
XX AC ABBI1470;
XX AC ABBI1470;
DT 11-JAN-2002 (first entry)
XX Human Huntington's disease protein homologue, SEQ ID NO:1840.
KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW inhibit; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW chronic inflammatory condition; proliferative retinopathy;
KW atherosclerosis; coronary heart disease; arterial ischaemia;

bone disorder; osteoporosis; vascular growth disorder;
tissue regeneration; wound healing; infection; immune disorder;
cell culture; drug screening; gene therapy; antiinflammatory;
antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
antifungal; vulnery; antiulcer.
Homo sapiens.
WO200157188-A2.
09-AUG-2001.
05-FEB-2001; 2001WO-US03800.
03-FEB-2000; 2000US-0496914.
27-APR-2000; 2000US-0560875.
(HYSE-) HYSEQ INC.
Tang YT, Liu C, Drmanac RT;
WPI; 2001-457740/49.
N-PSDB; ABA08714.
Human proteins and DNA encoding sequences useful for preventing,
treating or ameliorating a medical condition in a mammalian subject
e.g. arthritis and cancer -
Claim 20; Page 195-196; 1963pp; English.
Sequences ABBI0981-ABBI2330 represent 1350 novel human polypeptides, and
sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
invention also relates to vectors and recombinant host cells comprising a
nucleotide of the invention, methods of producing the novel polypeptides,
antibodies against the polypeptides, methods of detecting the nucleotides
or polypeptides in a sample, and methods of identifying compounds which
bind to polypeptides of the invention. Although novel, many of the
polypeptides of the invention have homology to known proteins, thereby
giving an insight into their probable biological activities, and hence
potential therapeutic applications. The polypeptides of the invention may
have various activities, including cytokine, cell proliferation or cell
differentiation activities; stem cell growth factor activity;
haematopoiesis regulatory activity; tissue growth activity;
immunomodulatory activity; activin or inhibin-related activities;
chemotactic or chemokinetic activities; haemostatic, thrombotic or
thrombolytic activities; receptor or ligand activities; or may be
involved in oncogenesis, cancer cell proliferation or metastasis.
Depending on their biological activities, polypeptides and nucleotides of
the invention are useful for preventing, treating or ameliorating medical
conditions, e.g., by protein or gene therapy. Such conditions include
cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
proliferative retinopathy, atherosclerosis, coronary heart disease,
arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
vascular growth. Polypeptides involved with tissue regeneration and
repair (or nucleic acids encoding them) may be used to promote wound
healing (e.g., of burns, incisions and ulcers), while those with
immunomodulatory activities may be used in the treatment of viral,
bacterial and fungal infections in addition to immune disorders.
Polypeptides with growth factor activity may be used in cell cultures to
promote cell growth. For example, such polypeptides may be used to
manipulate stem cells in culture to give rise to neuroepithelial cells
that can be used to augment or replace cells damaged by illness,
autoimmune disease or accidental damage. The polypeptides and nucleotides
may also be used in the diagnosis of the above conditions, and in drug
screening techniques. The present sequence represents a novel human
polypeptide of the invention.

Sequence 3223 AA;

Query Match

99.1%; Score 7821.5; DB 22; Length 3223;

Best Local Similarity 99.8%; Pred. No. 0;

[illegible]

QY	1078	FPLDLSAHQDALILAGNLLAASAPKSLKRSWASEEENPAATKQEEVWPALGORALVPMV	11337
Db	1160	FPLDLSAHQDALILAGNLLAASAPKSLKRSWASEEENPAATKQEEVWPALGORALVPMV	1219
QY	1138	EQLFSLHLKVINICARHVLDDVAPGPAIKAAALPSLTNPPLSPIRRCKEKEPEGEASVPL	1197
Db	1220	EQLFSLHLKVINICARHVLDDVAPGPAIKAAALPSLTNPPLSPIRRCKEKEPEGEASVPL	1279
QY	1198	SPKKGSEASAAASRQSDTSGPVTTKSSSGLGSFYHLPSVLKLRHDLVKATHANYKVTLDLQN	1257
Db	1280	SPKKGSEASAAASRQSDTSGPVTTKSSSGLGSFYHLPSVLKLRHDLVKATHANYKVTLDLQN	1339
QY	1258	STEKFGGFLRSALDVLVSQILELATLQDIGKCVBEELGYLKSFCFSREPMMATVCVQQLLKT	1317
Db	1340	STEKFGGFLRSALDVLVSQILELATLQDIGKCVBEELGYLKSFCFSREPMMATVCVQQLLKT	1399
QY	1318	LFGTNLASQFDGLSSNPSSQSGRAQRLGSSSVRPLCYHYCFMAYPTHFTQALADASLRNM	1377
Db	1400	LFGTNLASQFDGLSSNPSSQSGRAQRLGSSSVRPLCYHYCFMAYPTHFTQALADASLRNM	1459
QY	1378	VQAEQENDTSGWFDVLQKVSTQLTNLTSVTKNKRADKNAIHNHRLFEPLVIRKALKOYTT	1437
Db	1460	VQAEQENDTSGWFDVLQKVSTQLTNLTSVTKNKRADKNAIHNHRLFEPLVIRKALKOYTT	1519
QY	1438	TTCVQLQKQVLDLLAQLVOLRVNVCLLDSQVFTGFVLKQFEYIEVQGFRESEAIIPNIF	1497
Db	1520	TTCVQLQKQVLDLLAQLVOLRVNVCLLDSQVFTGFVLKQFEYIEVQGFRESEAIIPNIF	1579
QY	1498	FFLVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1533	
Db	1580	FFLVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1615	
RESULT 5			
AAR58777			
ID	AAR58777 standard; Protein; 3144 AA.		
XX	AAR58777;		
XX	13-APR-1995 (first entry)		
DT	Protein encoded by Huntingtin DNA\IT15 gene.		
XX	Polymerase chain reaction; primer: PCR; amplify; Huntingtin; IT15;		
KW	open reading frame; polymorphic; (CAG)n; trinucleotide repeat;		
KW	allele; Huntingdon's disease; HD; chromosome; juvenile HD; unstable;		
KW	expandable; linkage disequilibrium.		
XX	Homo sapiens.		
XX	EP614977-A.		
XX	14-SEP-1994.		
XX	07-MAR-1994; 94EP-0301587.		
XX	05-MAR-1993; 93US-0027498.		
PR	01-JUL-1993; 93US-0085000.		
XX	(GEO) GEN HOSPITAL CORP.		
PA	Ambrose CM, Duyao MP, Gusella JF, MacDonald ME;		
XX	WPI; 1994-281205/35.		
XX	N-PSDB; AAQ67401.		
XX	New huntingtin protein and related nucleic acid, antibodies etc.		
PT	- for treatment and diagnosis of neuro-degenerative disease,		
PT	specifically Huntington's disease		
XX	Claim 1; Page 22-33; 66pp; English.		
PS			
XX			

XX 05-MAR-1993; 93US-0027498.
 PR 01-JUL-1993; 93US-0085000.
 XX (GEO) GEN HOSPITAL CORP.
 PA Ambrose CM, Duyao WP, Gusella JF, MacDonald ME;
 XX WPI; 1997-558144/51.
 DR N-PSDB; AAT97924.
 XX Nucleic acid encoding huntingtin protein - useful for gene therapy
 PT of Huntington's disease
 XX Claim 2; Fig 4; 112pp; English.
 XX The present sequence represents a previously undescribed protein,
 CC encoded by a novel gene, termed huntingtin or htt. The huntingtin
 CC reading frame contains a polymorphic (CAG)_n trinucleotide repeat with at
 CC least 17 alleles in the normal population, varying from about 11 to 34
 CC CAG copies. Huntington's disease (HD) is a progressive neurodegenerative
 CC disorder characterised by motor disturbance, cognitive loss and
 CC psychiatric manifestations. The genetic defect causing HD is assigned to
 CC chromosome 4. On HD chromosomes, the length of the trinucleotide CAG
 CC repeat is substantially increased, e.g. about 37 to at least 73 copies.
 CC The huntingtin gene and proteins encoded by it, may be used for the
 CC diagnosis or treatment of Huntington's disease. The huntingtin gene
 CC is especially used in gene therapy of a symptomatic or presymptomatic
 CC patient. The method comprises providing a functional huntingtin gene with
 CC a (CAG)_n repeat of the normal range of 11-34 copies, or an antisense
 CC sequence, to the desired cells of the patient, in a manner that permits
 CC the expression of the huntingtin protein provided by the gene, or
 CC inhibits expression of the mutated huntingtin gene, for a time and in a
 CC quantity sufficient to provide the huntingtin function to the cells of
 CC the patient.
 XX Sequence 3144 AA:

Query Match 98.9%; Score 7805.5; DB 18; Length 3144;
 Best Local Similarity 99.6%; Pred. No. 0;
 Matches 1530; Conservative 2; Mismatches 1; Indels 3; Gaps 1:

QY 1 MATLEKIMAFESLKSF---QQQQQQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPQA 57
 DB 1 MATLEKIMAFESLKSFQQQQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPQA 60
 QY 58 QPLLPPQPP 117
 DB 61 QPLLPPQPP 120
 QY 118 PEFQKLLGIAMFLFLLCSDDAESDVRWVADECLNKVIKALMDSNLPRLQELYLEYKEIKNG 177
 DB 121 PEFQKLLGIAMFLFLLCSDDAESDVRWVADECLNKVIKALMDSNLPRLQELYLEYKEIKNG 180
 QY 178 APRSLRAALWRFELAEHLVRPQKRPYLVNLLPCLPRTSKRPESVQETLAAAVPKIMAS 237
 DB 181 APRSLRAALWRFELAEHLVRPQKRPYLVNLLPCLPRTSKRPESVQETLAAAVPKIMAS 240
 QY 238 FGNFANDNEIKVLLKAFIANLKSSPTIRTAAGSAVSIQHSRRRTQYFYSWLLNVLLGL 297
 DB 241 FGNFANDNEIKVLLKAFIANLKSSPTIRTAAGSAVSIQHSRRRTQYFYSWLLNVLLGL 300
 QY 298 LVPVEDHSTLLILGVLLTLYLVPLLOQQVDTSLKSGFVTRKEMEVSPSAEQLVQVY 357
 DB 301 LVPVEDHSTLLILGVLLTLYLVPLLOQQVDTSLKSGFVTRKEMEVSPSAEQLVQVY 360
 QY 358 ELTLHHTQHODHNVVGTGALELLQQLFRTPPPELLQTLTAVGGTGLTAAKESGGRSRG 417
 DB 361 ELTLHHTQHODHNVVGTGALELLQQLFRTPPPELLQTLTAVGGTGLTAAKESGGRSRG 420
 QY 418 SIVELIAGGGSSCPVLSRKQKGVLLGEEALEDDSESDSVSSSALTATSVKDEISGEL 477
 DB 421 SIVELIAGGGSSCPVLSRKQKGVLLGEEALEDDSESDSVSSSALTATSVKDEISGEL 480

QY 478 AASSGVSTPGSAGHDIITEQPSOHTLQADSVDLASCDLTSSATDGDDEEDILSHSSQVS 537
 DB 481 AASSGVSTPGSAGHDIITEQPSOHTLQADSVDLASCDLTSSATDGDDEEDILSHSSQVS 540
 QY 538 AVPSDPAAMDNDGTQASSPISDSQTTTEGPDSAVTPSDSSEIVLDGTNQYLGILQIQP 597
 DB 541 AVPSDPAAMDNDGTQASSPISDSQTTTEGPDSAVTPSDSSEIVLDGTNQYLGILQIQP 600
 QY 598 QDEDEEATGILPDEASEAFRNSMALQOAHLLKNMHCROPSSDSSVDKFLVLRDEATEPGD 657
 DB 601 QDEDEEATGILPDEASEAFRNSMALQOAHLLKNMHCROPSSDSSVDKFLVLRDEATEPGD 660
 QY 658 QENKPCRIRKDGIGOSTDDDSAPLVHCVRLLSASFLLTGKGNVLVDPDRDVRVSVKALALSC 717
 DB 661 QENKPCRIRKDGIGOSTDDDSAPLVHCVRLLSASFLLTGKGNVLVDPDRDVRVSVKALALSC 720
 QY 718 VGAVALHPESFFSKLYKVPDLTTEYPEEQVYSDILNYIDHGDPOVRGATAILCGTGLICS 777
 DB 721 VGAVALHPESFFSKLYKVPDLTTEYPEEQVYSDILNYIDHGDPOVRGATAILCGTGLICS 780
 QY 778 ILSRSRHFVGDWMCITRTLTGNTFSLADCIPLLRKTLKDESSVTCKLACTAVRNCVMSLC 837
 DB 781 ILSRSRHFVGDWMCITRTLTGNTFSLADCIPLLRKTLKDESSVTCKLACTAVRNCVMSLC 840
 QY 838 SSSYSELGLQLIIDVTLNRSSYWLVRTELETLAEIDFLVSFLEAKAENLHRGAHHYT 897
 DB 841 SSSYSELGLQLIIDVTLNRSSYWLVRTELETLAEIDFLVSFLEAKAENLHRGAHHYT 900
 QY 898 GLLKLQERVLNNVHLLGDEDDPRVHVAASLRVLPKLFYKCDQOQADPVAVARDQS 957
 DB 901 GLLKLQERVLNNVHLLGDEDDPRVHVAASLRVLPKLFYKCDQOQADPVAVARDQS 960
 QY 958 SVYLKLLMHETQPPSHFSVSITRIYRGYNLLPSITDVTMNNLSRVIAAVSHELITSTT 1017
 DB 961 SVYLKLLMHETQPPSHFSVSITRIYRGYNLLPSITDVTMNNLSRVIAAVSHELITSTT 1020
 QY 1018 RALTFGCCEALCLLSTAFPVCIWSLGHGCVPPPLSASDESRSKCTVGMMATMILTLSSAW 1077
 DB 1021 RALTFGCCEALCLLSTAFPVCIWSLGHGCVPPPLSASDESRSKCTVGMMATMILTLSSAW 1080
 QY 1078 FPLDLSAHQDALILAGNLLAASAPKSLRSSWASBEENAPATKOEYVMPALGDRLALVPMV 1137
 DB 1081 FPLDLSAHQDALILAGNLLAASAPKSLRSSWASBEENAPATKOEYVMPALGDRLALVPMV 1140
 QY 1138 EQLFSLHLKLVINICAHVLDDVAPGPAIKAALPSTNPPSLPIRRKGEKEPGEQASVPL 1197
 DB 1141 EQLFSLHLKLVINICAHVLDDVAPGPAIKAALPSTNPPSLPIRRKGEKEPGEQASVPL 1200
 QY 1198 SPKKGSEASASRSQSDTSGPVTTTSSSLGSGFYHLPYSLKLHDVLKATHANYKVTLDLQN 1257
 DB 1201 SPKKGSEASASRSQSDTSGPVTTTSSSLGSGFYHLPYSLKLHDVLKATHANYKVTLDLQN 1260
 QY 1258 STEKEGGFLRSALDVLISOILELATLQDQKGVVEETLGLKSCFSREPMATVVCVQQLLKT 1317
 DB 1261 STEKEGGFLRSALDVLISOILELATLQDQKGVVEETLGLKSCFSREPMATVVCVQQLLKT 1320
 QY 1318 LFGTNLASQFDGLSSNPSPKSGRAQLGSSSVRPGLYHYCFMAPYTHFTQALADASLRNM 1377
 DB 1321 LFGTNLASQFDGLSSNPSPKSGRAQLGSSSVRPGLYHYCFMAPYTHFTQALADASLRNM 1380
 QY 1378 VQAEQENDTSGWFDVLQKVSTQLKTNLTSVTKNRADKNAIHNHILRFLPEPLVIKALKQYTT 1437
 DB 1381 VQAEQENDTSGWFDVLQKVSTQLKTNLTSVTKNRADKNAIHNHILRFLPEPLVIKALKQYTT 1440
 QY 1438 TTCVQLQKQVLDLLAQVLQVRVNYCLLSDQVFIGFVLKQFEYIEVGQFRESEALIPNIF 1497
 DB 1441 TTCVQLQKQVLDLLAQVLQVRVNYCLLSDQVFIGFVLKQFEYIEVGQFRESEALIPNIF 1500
 QY 1498 FFLVLVLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1533
 DB 1501 FFLVLVLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1536

Db 714 ALALSCVGAVALHPSEFFSKLYKAPLDSVEYQDEQVSDILSYIHGGDPQVRGATAILC 773
QY 772 GTLICSLSRPHVGDWGTIRLTGNTFSLADCTPILRLKTLKDESSVTCKLACTAVRN 831
Db 774 GTLVLSILSRPHVGDWGTIRLTGNTFSLADCTPILRLKTLKDESSVTCKLACTAVRL 833
QY 832 CVNSLGSYSSEIGLQIILDVLTNRSSYWLVRTELLETAEIDFRLVSFLKAKAENLHR 891
Db 834 CVAALCGSSYSOWGLQILITDLTLRSSYWLVRTELLETAEIDFRLVSFLKAKAENLHR 893
QY 892 GAHHYTGLLKQRLVNLNNVIHLLGDEDPVRHVAASLRLVPKLFYKCDQGQADPVVA 951
Db 894 GAHHYTGLLKQRLVNLNNVIHLLGDEDPVRHVAASLRLVPKLFYKCDQGQADPVVA 953
QY 952 VARDQSSVYLKILMHETQPPSHFSVSTIRYRGYNLPLSDITDMENNLRSVIAAVSHE 1011
Db 954 VARDQSSVYLKILMHETQPPSHFSVSTIRYRGYNLPLSDITDMENNLRSVIAAVSHE 1013
QY 1012 LITSTTRALTFCGCEALCLLSTAFPCVWSLGHWCVPPLSGAASDESKSCVTGMMATMIT 1071
Db 1014 LITSTTRALTFCGCEALCLLSTAFPCVWSLGHWCVPPLSGAASDESKSCVTGMMATMIT 1073
QY 1072 LLSAWFPLDLAHOADALILAGNLAASAPKSLRSWASEEPAANPAATQKEEVPALGDR 1131
Db 1074 LLSAWFPLDLAHOADALILAGNLAASAPKSLRSWASEEPAANPAATQKEEVPALGDR 1133
QY 1132 ALVPVVEQLFSLHLLKVINICAHVLDVAPGPAKALPSLTNPSPSLSPTRRKGKEPGE 1191
Db 1134 SLVPVVEQLFSLHLLKVINICAHVLDVAPGPAKALPSLTNPSPSLSPTRRKGKEPGE 1193
QY 1192 QASVPLSPKKGEASAAARSQSDTSGPTVTSKSSLSGFSYHLPYSLKLDHVLKATHANKV 1251
Db 1194 QASVPLSPKKGEASAAARSQSDTSGPTVTSKSSLSGFSYHLPYSLKLDHVLKATHANKV 1253
QY 1252 TLDLQNSTEKGFEGLRSALDVLQSLLELATLQDICKCYEEILGYLKSCFSREPMMATVCV 1311
Db 1254 TLDLQNSTEKGFEGLRSALDVLQSLLELATLQDICKCYEEILGYLKSCFSREPMMATVCV 1313
QY 1312 QOLLKTLFTNLASQFDGLSNPSKSGRAQRLGSSVVRPGLYHYCFMAYTHFTQALAD 1371
Db 1314 QOLLKTLFTNLASQFDGLSNPSKSGRAQRLGSSVVRPGLYHYCFMAYTHFTQALAD 1373
QY 1372 ASLRNVAQENDTSGHFDVLQKYSTQKTNLTSTVTKNRADKNAIHNRHIFLPEPLVKA 1431
Db 1374 ASLRNVAQENDTSGHFDVLQKYSTQKTNLTSTVTKNRADKNAIHNRHIFLPEPLVKA 1433
QY 1432 LKQYTTTTVCVQLOKQVLDLALQVLRLVNYCLLDSQVDFIGFVLKQFEYIEVGQFRESEA 1491
Db 1434 LKQYTTTTVCVQLOKQVLDLALQVLRLVNYCLLDSQVDFIGFVLKQFEYIEVGQFRESEA 1493
QY 1492 IIPNIFLVLVLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1533
Db 1494 IIPNIFLVLVLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1535

RESULT 11
AA133503
ID AAY33503 standard; Protein: 589 AA.
XX
AC AAY33503;
XX
XX
DT 19-JAN-2000 (first entry)
XX
DE Human apopain cleavage fragment 3.
XX
KW Proapoptotic; dependence domain; p75NTR; androgen receptor; DCC;
KW huntingtin polypeptide; Machado-Joseph disease; SCAL; SCA2; SCA6;
KW atrophin-1; cell death; apoptosis; Huntington's disease; head trauma;
KW Alzheimer's disease; Kennedy's disease; spinocerebellar ataxia; stroke;
KW dentatorubropallidoluysian atrophy; cell proliferation; cell survival;
KW neoplastic; malignant; autoimmune; fibrotic; apopain.
OS Homo sapiens.

XX WO9945944-A1.
PN 16-SEP-1999.
XX
PD 11-MAR-1999; 99WO-US05250.
XX
PR 12-MAR-1998; 98US-0041886.
XX
PA (BURN-) BURNHAM INST.
XX
PI Bredesen DE, Rabizadeh S;
XX
XX WPI; 1999-561617/47.
XX
PT New proapoptotic dependence peptides, used to develop products for
PT treating, e.g. Alzheimer's disease -
XX
XX Disclosure; Page 176-178; 1999p; English.
XX
PS This invention describes novel pure proapoptotic dependence peptides
CC which comprise a sequence of an active dependence domain selected from
CC dependence polypeptides consisting of p75NTR, androgen receptor, DCC,
CC huntingtin polypeptide, Machado-Joseph disease gene product, SCAL, SCA2,
CC SCA6 and atrophin-1 polypeptide. The proapoptotic peptides are capable
CC of inducing cell death and can be used to develop products to mediate or
CC inhibit apoptosis. The methods can be used for reducing the severity of
CC a proapoptotic dependence domain mediated pathological conditions e.g.
CC Huntington's disease, Alzheimer's disease, Kennedy's disease,
CC Spinocerebellar ataxia, dentatorubropallidoluysian atrophy, be used for
CC Machado-Joseph disease, stroke or head trauma. They can also be used for
CC reducing the severity of a pathological condition mediated by upregulated
CC cell proliferation or cell survival e.g. neoplastic, malignant,
CC autoimmune or fibrotic conditions. This sequence represents a human
CC apopain cleavage fragment described in the method of the invention.
XX
SQ Sequence 589 AA;
Query Match 37.6%; Score 2967.5; DB 20; Length 589;
Best Local Similarity 99.5%; Pred. No. 4.2e-205;
Matches 586; Conservative 0; Mismatches 0; Indels 3; Gaps 1;
QY 1 MATLEKLMKAFESLKSF---QQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ 57
Db 1 MATLEKLMKAFESLKSFQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ 60
QY 58 QPLLQPPQPP 117
Db 61 QPLLQPPQPP 120
QY 118 PEFQKLLGIAMELFLLCSDAEDSVRMVADECLNKVKIKALMDSNLPRLQLELYKEIKKNG 177
Db 121 PEFQKLLGIAMELFLLCSDAEDSVRMVADECLNKVKIKALMDSNLPRLQLELYKEIKKNG 180
QY 178 APRSLRAALWRFALHVRPQKCRPYLVNLLPCLTRTSKRPEESVOETLAAAVPKTMS 237
Db 181 APRSLRAALWRFALHVRPQKCRPYLVNLLPCLTRTSKRPEESVOETLAAAVPKTMS 240
QY 238 FGNFANDNETKVLKAFIANLKSSSPTRRTAAGSAVSICQHSRRTOYFYSWLLNVLGL 297
Db 241 FGNFANDNETKVLKAFIANLKSSSPTRRTAAGSAVSICQHSRRTOYFYSWLLNVLGL 300
QY 298 LVPVEDHSTLLTILGVLLTLRYLPLLQOQVKTSLKSGFVTRKMEVSPSAEQLVQVY 357
Db 301 LVPVEDHSTLLTILGVLLTLRYLPLLQOQVKTSLKSGFVTRKMEVSPSAEQLVQVY 360
QY 358 ELTLHHTQHODHNWVTALELLQQLFRTPELQTLTAVGGIGQLTAAKEESGGRSRG 417
Db 361 ELTLHHTQHODHNWVTALELLQQLFRTPELQTLTAVGGIGQLTAAKEESGGRSRG 420
QY 418 SIVELIAGGSSCSPVLSRRKQKGVLLGEEALEDDSESRSVSSSALTASVKDEISGEL 477
Db 421 SIVELIAGGSSCSPVLSRRKQKGVLLGEEALEDDSESRSVSSSALTASVKDEISGEL 480

XX PS Disclosure; Page 173-174; 199pp; English.

XX CC This invention describes novel pure proapoptotic dependence peptides

CC CC which comprise a sequence of an active dependence domain selected from

CC CC dependence polypeptides consisting of p75NTR, androgen receptor, DCC,

CC CC huntingtin polypeptide, Machado-Joseph disease gene product, SCAl, SCA2,

CC CC SCA6 and atrophin-1 polypeptide. The proapoptotic peptides are capable

CC CC of inducing cell death and can be used to develop products to mediate or

CC CC inhibit apoptosis. The methods can be used for reducing the severity of

CC CC a proapoptotic dependence domain mediated pathological conditions e.g.

CC CC Huntington's disease, Alzheimer's disease, Kennedy's disease,

CC CC Spinocerebellar ataxias, dentatorubropallidoluysian atrophy,

CC CC Machado-Joseph disease, stroke or head trauma. They can also be used for

CC CC reducing the severity of a pathological condition mediated by upregulated

CC CC cell proliferation or cell survival e.g. neoplastic, malignant,

CC CC autoimmune or fibrotic conditions. This sequence represents a human

CC CC apocain cleavage fragment described in the method of the invention.

XX SQ Sequence 530 AA;

Query Match 33.9%; Score 2676.5; DB 20; Length 530;
 Best Local Similarity 99.4%; Pred. No. 3.5e-184;
 Matches 527; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 1 MATLEKLMKAFESLKSF---QQQQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPQ 57
 DB 1 MATLEKLMKAFESLKSFOQQQQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPQ 60

QY 58 QPLLPOPP 117
 DB 61 QPLLPOPP 120

QY 118 PEQKLLGAMELFLCSDDAESDVRVADECLNKVKIKALMDSNLPRLQLEYKEIKNG 177
 DB 121 PEQKLLGAMELFLCSDDAESDVRVADECLNKVKIKALMDSNLPRLQLEYKEIKNG 180

QY 178 APRSLRAALWRFALHVRPQKCRPYLVNLLPCLTRTSKRPEESVQETLAAAVPKIMAS 237
 DB 181 APRSLRAALWRFALHVRPQKCRPYLVNLLPCLTRTSKRPEESVQETLAAAVPKIMAS 240

QY 238 FGNFANDNEIKVLLKAFIANLKSSSPTRRTAAGSAVSTICQHSRRTQYFYSWLLNVLLGL 297
 DB 241 FGNFANDNEIKVLLKAFIANLKSSSPTRRTAAGSAVSTICQHSRRTQYFYSWLLNVLLGL 300

QY 298 LVPVEDEHSTLLILGLVLLTLRYLVPLLQQQVKDTSLKSGFVTRKEMEVSFAEQLVQVY 357
 DB 301 LVPVEDEHSTLLILGLVLLTLRYLVPLLQQQVKDTSLKSGFVTRKEMEVSFAEQLVQVY 360

QY 358 ELTLHHTQHODHNVTGALQLLOQLFRTPPPELLQTLTAVGGIGQITAAKESGGRSRG 417
 DB 361 ELTLHHTQHODHNVTGALQLLOQLFRTPPPELLQTLTAVGGIGQITAAKESGGRSRG 420

QY 418 SIVELIAGGSGSPVLSRKQKGVLLGEEALEDDSDRSVSSSALTASVKDEISGEL 477
 DB 421 SIVELIAGGSGSPVLSRKQKGVLLGEEALEDDSDRSVSSSALTASVKDEISGEL 480

QY 478 AASSGVSTPGSAGHDITTEQPSRQHTLQADSVDLASCDLTSSATGDDEED 527
 DB 481 AASSGVSTPGSAGHDITTEQPSRQHTLQADSVDLASCDLTSSATGDDEED 530

RESULT 14
 AAY33500
 ID AAY33500 standard; Protein; 513 AA.
 XX AC AAY33500;
 XX AC
 DT 19-JAN-2000 (first entry)
 XX Human huntingtin polypeptide amino terminal fragment.
 DE Proapoptotic; dependence domain; p75NTR; androgen receptor; DCC;
 KW

KW huntingtin polypeptide; Machado-Joseph disease; SCAl; SCA2; SCA6;
 KW atrophin-1; cell death; apoptosis; Huntington's disease; head trauma;
 KW Alzheimer's disease; Kennedy's disease; spinocerebellar ataxia; stroke;
 KW dentatorubropallidoluysian atrophy; cell proliferation; cell survival;
 KW neoplastic; malignant; autoimmune; fibrotic.

OS Homo sapiens.

PN WO9945944-A1.

XX 16-SEP-1999.

XX 11-MAR-1999; 99WO-US05250.

PR 12-MAR-1998; 98US-0041886.

PA (BURN-) BURNHAM INST.

XX Bredesen DE, Rabizadeh S;

XX WPI; 1999-561617/47.

DR New proapoptotic dependence peptides, used to develop products for
 treating, e.g. Alzheimer's disease -

PT Disclosure; Page 171-172; 199pp; English.

PS This invention describes novel pure proapoptotic dependence peptides
 which comprise a sequence of an active dependence domain selected from
 dependence polypeptides consisting of p75NTR, androgen receptor, DCC,
 huntingtin polypeptide, Machado-Joseph disease gene product, SCAl, SCA2,
 SCA6 and atrophin-1 polypeptide. The proapoptotic peptides are capable
 of inducing cell death and can be used to develop products to mediate or
 inhibit apoptosis. The methods can be used for reducing the severity of
 a proapoptotic dependence domain mediated pathological conditions e.g.
 Huntington's disease, Alzheimer's disease, Kennedy's disease,
 Spinocerebellar ataxias, dentatorubropallidoluysian atrophy,
 Machado-Joseph disease, stroke or head trauma. They can also be used for
 reducing the severity of a pathological condition mediated by upregulated
 cell proliferation or cell survival e.g. neoplastic, malignant,
 autoimmune or fibrotic conditions. This sequence represents the human
 huntingtin polypeptide amino terminal fragment described in the method of
 the invention.

XX SQ Sequence 513 AA;

Query Match 32.8%; Score 2589.5; DB 20; Length 513;
 Best Local Similarity 99.4%; Pred. No. 6.2e-178;
 Matches 510; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 1 MATLEKLMKAFESLKSF---QQQQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPQ 57
 DB 1 MATLEKLMKAFESLKSFOQQQQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPQ 60

QY 58 QPLLPOPP 117
 DB 61 QPLLPOPP 120

QY 118 PEQKLLGAMELFLCSDDAESDVRVADECLNKVKIKALMDSNLPRLQLEYKEIKNG 177
 DB 121 PEQKLLGAMELFLCSDDAESDVRVADECLNKVKIKALMDSNLPRLQLEYKEIKNG 180

QY 178 APRSLRAALWRFALHVRPQKCRPYLVNLLPCLTRTSKRPEESVQETLAAAVPKIMAS 237
 DB 181 APRSLRAALWRFALHVRPQKCRPYLVNLLPCLTRTSKRPEESVQETLAAAVPKIMAS 240

QY 238 FGNFANDNEIKVLLKAFIANLKSSSPTRRTAAGSAVSTICQHSRRTQYFYSWLLNVLLGL 297
 DB 241 FGNFANDNEIKVLLKAFIANLKSSSPTRRTAAGSAVSTICQHSRRTQYFYSWLLNVLLGL 300

QY 298 LVPVEDEHSTLLILGLVLLTLRYLVPLLQQQVKDTSLKSGFVTRKEMEVSFAEQLVQVY 357
 DB 301 LVPVEDEHSTLLILGLVLLTLRYLVPLLQQQVKDTSLKSGFVTRKEMEVSFAEQLVQVY 360


```
QY 241 FANDNEIKVLLKAFIANLKSSPTIRRTAAGSAVSTCOHSRSTOYFYSWLLNVLLGLLVP 300
Db 241 FANDNEIKVLLKAFIANLKSSPTIRRTAAGSAVSTCOHSRSTOYFYSWLLNVLLGLLVP 300
QY 301 VEDEHSTLILGLVLLTRVLYVLLQOVKDTSLKGSFGVTRKEMEVSPSAEQLVQVYELT 360
Db 301 VEDEHSTLILGLVLLTRVLYVLLQOVKDTSLKGSFGVTRKEMEVSPSAEQLVQVYELT 360
QY 361 LHHTQHODHNVTGALVLELLOQLFRTPPPPELLQTLTAVGGIGOLTAAKESGGRSGSIV 420
Db 361 LHHTQHODHNVTGALVLELLOQLFRTPPPPELLQTLTAVGGIGOLTAAKESGGRSGSIV 420
QY 421 ELIAGGSCSPVLSRKQKGVLLGEEALEDDSESRDVSSSALTASVKDISGELAAS 480
Db 421 ELIAGGSCSPVLSRKQKGVLLGEEALEDDSESRDVSSSALTASVKDISGELAAS 480
QY 481 SGVSTPGSAGHDITRQPSQHTLOADSVLDLASCIDLTSATDGEDILSHSSQSVAVP 540
Db 481 SGVSTPGSAGHDITRQPSQHTLOADSVLDLASCIDLTSATDGEDILSHSSQSVAVP 540
QY 541 SDPAMLDNGTQASSPISDSSTTTTGGPSAVTPSDSSEIVLDGTDNQVLGIGQPQDE 600
Db 541 SDPAMLDNGTQASSPISDSSTTTTGGPSAVTPSDSSEIVLDGTDNQVLGIGQPQDE 600
QY 601 DEBATGLPDEASEAFRNSMALQQAHLKNNKSHCROPDSSVDKFLVLRDEATEPGDQEN 660
Db 601 DEBATGLPDEASEAFRNSMALQQAHLKNNKSHCROPDSSVDKFLVLRDEATEPGDQEN 660
QY 661 KPCRIGDQSDTDDPSAPLVHCVRLLSASFLLTGKKNVLPDRDVRVSVKALASCVGA 720
Db 661 KPCRIGDQSDTDDPSAPLVHCVRLLSASFLLTGKKNVLPDRDVRVSVKALASCVGA 720
QY 721 AVALHPESFSKLYKVPDLTTEPEQYVSDIILNYIDHGDQVQRGATAILCGTLICSIIS 780
Db 721 AVALHPESFSKLYKVPDLTTEPEQYVSDIILNYIDHGDQVQRGATAILCGTLICSIIS 780
QY 781 RSRFHVGDMWGTIRTLTGNTFSLADCIPLRLKTLKDESSVTCKLACTAVRNCVMSLCSSS 840
Db 781 RSRFHVGDMWGTIRTLTGNTFSLADCIPLRLKTLKDESSVTCKLACTAVRNCVMSLCSSS 840
QY 841 YSELGLQLIIDVLTNRSSVYLVRTLELLETAEIDFRLVFLSEAKAENLRGAAHHTTGLL 900
Db 841 YSELGLQLIIDVLTNRSSVYLVRTLELLETAEIDFRLVFLSEAKAENLRGAAHHTTGLL 900
QY 901 KLQERVLNNVVIHLGDEDPVRVHVAASLIRLVPLFKYKCDQOGADPVVAVARDQSSVY 960
Db 901 KLQERVLNNVVIHLGDEDPVRVHVAASLIRLVPLFKYKCDQOGADPVVAVARDQSSVY 960
QY 961 LKLLMHETQPPSHFSVSTIIRYGYNLLPSITDVTMNNLSRVIAAVSHELITSTTRAL 1020
Db 961 LKLLMHETQPPSHFSVSTIIRYGYNLLPSITDVTMNNLSRVIAAVSHELITSTTRAL 1020
QY 1021 TFGCCALCLLSTAFFVCIVNSLGHHCVPPLLSASDESRSKCTVGMATMILTLLSSAWFPL 1080
Db 1021 TFGCCALCLLSTAFFVCIVNSLGHHCVPPLLSASDESRSKCTVGMATMILTLLSSAWFPL 1080
QY 1081 DLSAQDALILAGNLLAASAPKSLRSSWASSEEPANPAATKQEEVWPALGDRALVPVVEQL 1140
Db 1081 DLSAQDALILAGNLLAASAPKSLRSSWASSEEPANPAATKQEEVWPALGDRALVPVVEQL 1140
QY 1141 FSHLLKVINICAHVLDVAVCPAIPAALPSITNPPSLPIRRKGKEKEPGEQASVPLSPK 1200
Db 1141 FSHLLKVINICAHVLDVAVCPAIPAALPSITNPPSLPIRRKGKEKEPGEQASVPLSPK 1200
QY 1201 KGSEASASROSDTSGPVTTSSKSSLSGFYHLPSYKLHDLVLAATHANYKVTLDLQNSTE 1260
Db 1201 KGSEASASROSDTSGPVTTSSKSSLSGFYHLPSYKLHDLVLAATHANYKVTLDLQNSTE 1260
QY 1261 KFGGFLRSALDVLSQLLELATLQDIGKVCVEILGYLKSFSRPMATVVCVQQLLTLFG 1320
Db 1261 KFGGFLRSALDVLSQLLELATLQDIGKVCVEILGYLKSFSRPMATVVCVQQLLTLFG 1320
```

```
QY 1321 TNLASQFDGLSSNPSKSGQRAQLGSSSVRGLYHYCFMAYPTHFTQALADASLRNNVQA 1380
Db 1321 TNLASQFDGLSSNPSKSGQRAQLGSSSVRGLYHYCFMAYPTHFTQALADASLRNNVQA 1380
QY 1381 EOENDTSGWFDVLOKVSTOLKTNLTSTVKRPNADKNAIHNHRLFEPLVIALKQYTTTTC 1440
Db 1381 EOENDTSGWFDVLOKVSTOLKTNLTSTVKRPNADKNAIHNHRLFEPLVIALKQYTTTTC 1440
QY 1441 VOLOKQVLDLQAQLVQLRVNYCCLLDSQVFGFVGLKQFEVIEVQGFRESEAIPIPIFFL 1500
Db 1441 VOLOKQVLDLQAQLVQLRVNYCCLLDSQVFGFVGLKQFEVIEVQGFRESEAIPIPIFFL 1500
QY 1501 VLLSYERYHSHQIIIGIPKIIQOLCDGIMASGRKASQPQYRLCSP 1543
Db 1501 VLLSYERYHSHQIIIGIPKIIQOLCDGIMASGRKASQPQYRLCSP 1543
```

RESULT 2

```
US-10-042-417-32
; Sequence 32, Application US/10042417
; Patent No. US20020123082A1
; GENERAL INFORMATION:
; APPLICANT: Pagano, M.
; TITLE OF INVENTION: METHODS TO IDENTIFY COMPOUNDS USEFUL FOR THE TREATMENT OF
; TITLE OF INVENTION: PROLIFERATIVE AND DIFFERENTIATIVE DISORDERS
; FILE REFERENCE: 5914-090-999
; CURRENT APPLICATION NUMBER: US/10/042.417
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: 60/260,179
; PRIOR FILING DATE: 2001-01-5
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 32
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-042-417-32
```

Query Match 2.6%; Score 208; DB 12; Length 197;

Best Local Similarity 32.3%; Pred. No. 9.3e-06;

Matches 63; Conservative 12; Mismatches 60; Indels 60; Gaps 7;

```
QY 27 QQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ 77
Db 6 QQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ 62
QY 78 -----AAVEE-----PLHR-----PKKLSATKKDRVNHCLTICENI--VAQSV 114
Db 63 DDDVPADWVAEESGPGAQNSPYQLRRKTLTPKRTACTPKNSMEGASTSTTENFGHAKRA 122
QY 115 RNSPEFQKLLGIAMBELFLCSDDAESDVRMVADECLNKVIALMDSNL----- 162
Db 123 RVSGKSQDLSSAAPEQL-----OEKLPDEVVLKIFSYLLLEQDLCLRAACVCKRESE 173
QY 163 ----PRQLQELYKEI 173
Db 174 LANDPNLMKRLMEV 188
```

RESULT 3

US-09-770-689A-5

; Sequence 5, Application US/09770689A

; Patent No. US2002011517A1

; GENERAL INFORMATION:

; APPLICANT: YAN, Chunhua et al.

; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE

; TITLE OF INVENTION: PROTEINS, AND USES THEREOF

; FILE REFERENCE: CLO01079

; CURRENT APPLICATION NUMBER: US/09/770,689A

; CURRENT FILING DATE: 2001-01-29

; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: FastSeq for Windows Version 4.0

RESULT 7

Qy	608	----	LPDEASAFRNSMALQQA-	----	LLKNMSHCROPSSSVKFFVL-	640
Db	664	SVSLKKVSGTKEEIKEP-	ENINAALQETAPYISACOLIKETKLSAEPAPOFSYSEMA	722		
Qy	649	-----	RDEATEPGDQENKPKRIGD	ICQSDDD	SAPLV-----	HCVRLLSASFLLTG 695
Db	723	KVEQVPDHSSELVEDSP	SEPDVL-----	FSDDS	IPDVQKQDET	VWLKESLTETS 775
Qy	696	GKNVL-----	VPDRVRVSVKALSCVGA	VALHPE--	SFFSKLYKVP	LOTTEY 743
Db	776	FESMIEYENKEKLSALP	PEGGKPYLSEFSKLSLD	NTKDTLLP	DEVSTLSKKE	APIQMBEL 835
Qy	744	PEQYVSDIL	753			
Db	836	STAVYSNDDL	845			

RESULT 8

```

US-09-893-348-23
; Sequence 23, Application US/09893348
; Patent No. US20020072493A1
; GENERAL INFORMATION:
; APPLICANT: EISENBACH-SCHWARTZ, Michal
; APPLICANT: COHEN, Irun R.
; APPLICANT: BESERMAN, Pierre
; APPLICANT: MOSONEGO, Alon
; APPLICANT: MOALEM, Gila
; TITLE OF INVENTION: ACTIVATED T-CELLS, NERVOUS SYSTEM-SPECIFIC ANTIGENS AND THEIR
; FILE REFERENCE: EIS-SCHWARTZ=2A
; CURRENT APPLICATION NUMBER: US/09/893,348
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US 09/314,161
; PRIOR FILING DATE: 1999-05-19
; PRIOR APPLICATION NUMBER: US 09/218,277
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: PCT/US98/14715
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: IL 124500
; PRIOR FILING DATE: 1998-05-19
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 23
; LENGTH: 1192
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-893-348-23

```

Query Match 2.3%; Score 178.5; DB 10; Length 1192;
Best Local Similarity 19.4%; Pred. NO. 0.0067;
Matches 165; Conservative 114; Mismatches 302; Indels 269; Gaps 36;

Qy	34	QOQOQPPPPPPPPPPPPPPQOAPLQPQOQPPPPPPPPPPPPPPGPAVABEPHRRPKKELSA	93
Db	135	EDDEPPAPPPPPPPASV---SQEAPVMTTPAPAPAAPSTPA-----APKRRGSS	182
Qy	94	TKDVRVNHCLTTCENIVAOVSVRNSPEFOKLLIGIAEMLEFLICDDDAESVVRMADECLNKV	153
Db	183	GSVDETLFALPAASEPVIIRSSAENWDLKEQPGNTISA-----QOEDFPVSVLLETA---	232
Qy	154	IKALMDSNLPRLQLELYKEIKKNGAPRSURAAHLRFAELHLVRPKQCPYLVNLLPCLT	213
Db	233	-----ASUPSLS-----PLSAASFKEHE-----YGLNLSLTVP	260
Qy	214	RTSKRPEESVQETLAAAVPKIMASGFNFANDNEIKVLKAPTANLKSS---SPTRIRTA	270
Db	261	-----TEGTLQENVSASKEVSEKATLLIDRLTFESELYSEMGSSFSVSP-----KA	310
Qy	271	GSAVSICQHSRTOFYFYSHLLNVILGLLPVPEDEHSTLLILGVLLTLRYLVPLLQQOVKD	330
Db	311	ESAVIVA-----NPREIVKNKDEEEKLVSNMILHNOQEPLTALTUKE	356
Qy	331	TSL-----KGSFGVTRKMEVSPSAEQLVQVYEILTLLHHTQHODHNVVYTGALLELOOLFR	384

RESULT 7

```

US-09-789-386-2
; Sequence 2, Application US/09789386
; Patent No. US20020010324A1
; GENERAL INFORMATION:
; APPLICANT: MICHALOVICH, DAVID
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP-30165-C1
; CURRENT APPLICATION NUMBER: US/09/789,386
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: U.K. 9916898.1
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: U.K. 9816024.5
; PRIOR FILING DATE: 1998-07-22
; PRIOR APPLICATION NUMBER: US 09/359,208
; PRIOR FILING DATE: 1999-07-22
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1192
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-789-386-2

Query Match      2.3% Score 178.5; DB 10; Length 1192;
Best Local Similarity 19.4%; Pred. No 0.0067;
Matches 165; Conservative 114; Mismatches 302; Indels 269; Gaps 36;

QY   34 QQQQPPPPPPPPPPQLPQPQQPQAQLLPQPQQPPPPPPPPGPAVAEEPLHRPKKLSA 93
Db    : ||| ||||| : : : : : : : : : : : : : : : : : : : : : : : :
QY   135 EDEEPARPPPPPSV---SPOAEPTWTPAPAPAAPPSTEA-----APKRGS 182
Db    : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY   94 TKDRVNCLTICENIVAOSVRNSPEFKLIGTAMLEFLLCSDDAESDVRMWADCLNKV 153
Db    : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY   183 GSVDETFLALPAASEPIRIRSAENMDLKEQGWTISA-----QEEDPPSVILETA--- 232
Db    : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY   154 IKALMDSLNLRLQLELYKEIKNGAPSLRAALWRFAELAHLRVPQCRPYLVNLLPLCT 213
Db    : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY   233 -----ASLPSSL-----PLSAASFKEHE-----YLGNSLTVP 260
Db    : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY   214 RTSKRPEESVOETLAAPVKIMASFGNFANDNEIKVLKAFIANLKSS---SPTRRTAA 270
Db    : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY   261 -----TEGTGNSEAKSEKAKTLLDRLDTFESELEYSEMGSFVSVP-----KA 310
Db    : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY   271 GSAVICQHRSRTQFYFSWLNVLGLLVPEDEHSTLLILGVLLTRLVPLPQQOQVD 330
Db    : ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
QY   311 ESAVIVA-----NPREEIYNKDDEEKLVSNNILNQOELPTALTKLVE 356
Db    : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY   331 TSL-----KSGFGVTRKMEVSPAPQLVQVYELTLHHQTQHDHNVTGALELLQLFR 384
Db    : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY   357 DEWSSEKAKDSFNKKRVAVE-APMREEVADF----- 387
Db    : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY   385 TPPELLQTLTAVGGIGOLTAATKEEGSRSGSIIVELIAGGSCSPVLSRKQKGKVL 444
Db    : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY   388 KPFE-----RMVEVDKSIDS-----DMLAAGKKTIESNLESKVDK-KCFA 426
Db    : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY   445 G--EBEALDDSDRSVDSSALT-ASVKDEISGEL-----AASSGVST-----P 486
Db    : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY   427 DSLEQTNHKDSSNDOTSPFTPEGIDRPAGAYITCAPFPAAATESIATNIFFLLGDP 486
Db    : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY   487 GSAGH-----DIITEQPRSQHT-----LOADSVDLASCDLTSSA 520
Db    : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY   487 TSENKTDKKITEEKQAVITKNTSKTSNPFVLVAQDSETDYVTVDNLTKVTEEVANM 546
Db    : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY   521 TDGDEDILSHS-SQSYVASVPS-----DPAMDINDCTQASSPIDSS----SQTTTEGPDS 570
Db    : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY   547 PEGLTPLDLQVACESELNEVTGKIAYETKMDL---VQTSVMQESLYPAQLCPSEES 603
Db    : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY   571 AVTFSD-SSEVLDTGTONYL---GLQIQGDQDEDEATGI----- 607
Db    : ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
QY   604 EATSPVLPDITVMEAPLNSVAPSAGASVTQSSPSLEASSVYSIKHEPENPPPYEAM 663

```

```

Db 357 DEVVSEKAKDSFNERKVAE-APMREYADF----- 387
Qy 385 TPPELLQLTAVGGIGOLTAKEESGGRSRSIVELIAGGSSCPVLSRKQKGVLL 444
Db 388 -KPE-----RYVEKDKSDS-----DMLAAGKTESNLESKVDK-KCFA 426
Qy 445 G--EEEALEDSDSRSDVSSALT-ASVKDEISGEL-----AASSGVST-----P 486
Db 427 DSLEQINHEKDSSESSNDTSFSTPGIKDRPCAYITCAPFPAATESIATNIFPLGDP 486
Qy 487 GSAGH-----DIITEQPRSQHT-----LOADSVDLASCOTSSA 520
Db 487 TSENKTDKIEKKAQIVTEKNTSTKSNPFLVAAQDSETDYVTTDNLTKVTEEVVAMN 546
Qy 521 TDGDEEDILSHS--SSOVSAPVS-----DPAMDNDGTOASSPLSDS-----SOTTTEGPD 570
Db 547 PEGLTDLVQEAESSELENEVTGKIAYETKMDL---VQTEVQWESLYPAAQLCPSFEES 603
Qy 571 AVTPSD--SSEIVLDGTDNOYL---GLQIGQPODEDEATGI----- 607
Db 604 EATPSPVLPDIVEAPLNSAVPSAGASVIOPSSSPLEASSVNYESIKHEPNPPVEEAM 663
Qy 608 -----LPDASAFNRSSMALQOAH-----LLKNMSHCRQPSDSVVKFVL- 648
Db 664 SVSLKKVSGIKIEKEP-ENINAALOETEAPYISTACDLIKETKLSAEPAPDFSDYSEMA 722
Qy 649 -----RDEATEPGDQENKPCRKINGDIGQSDDDASPLV-----HCVRLLSASFLLTG 695
Db 723 KVEQVPDHSSELVEDSPSEVDL-----FSDSDIPIVQKQDETVMVKESLITETS 775
Qy 696 GKNVL-----VPDRDVRVSVKALSCVGAVALHPE--SPFSKLYKVPPLDTTEY 743
Db 776 FESMIEYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSPLSKKEIPLQMEEL 835
Qy 744 PEOQVSDIL 753
Db 836 STAVYSNDDL 845

RESULT 9
US-10-108-605-29
; Sequence 29, Application US/10108605
; Patent No. US20020160934A1
; GENERAL INFORMATION:
; APPLICANT: Broadus, Julie
; APPLICANT: Stam, Lynn
; APPLICANT: Bachmann, Jane
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT ENCODE
; FILE REFERENCE: 31133B
; CURRENT APPLICATION NUMBER: US/10/108,605
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: US 09/761,142
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/176,418
; PRIOR FILING DATE: 2000-01-14
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 29
; LENGTH: 1404
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-108-605-29

```

```

Query Match 2.2%; Score 177; DB 9; Length 1404;
Best Local Similarity 19.2%; Pred. No. 0.01;
Matches 170; Conservative 118; Mismatches 327; Indels 272; Gaps 37;

Qy 83 PLH-RPKKELSATKK-----DRVNHCLTICENIV-AQSVNSPFEQKLLGIAME 129
Db 307 PLHTRPTIYSVTAKVRGNPKALIIDIIVTQAGTYIKELVHGEFGRTPSLSSLIICKPMD 366

```

```

Qy 130 LFLCSDDAESDVRVADDECLNKV--IKALMDSNLPRLQLELYKEIKKN-----GAPRS 181
Db 367 IQALDVSTTESPAVHYISEDSESVTYVSGAVDCEV-----KDSNNQNSAVCAGAPLP 418
Qy 182 LRAALWRAEL-----AHLVRPKCRPYLVNLLPCLLRTSKR---PEESVOETL 227
Db 419 AIMIQRRAEAPIGAVQLTPLWEHILTRRL-PETIFPSAMYAEFHERLQDPQWQVQSHA 477
Qy 228 AAAPVKIMASFGNANDN-ETKVLKAFIANLKSSSTPIRTA-----AGSAVSIQ 278
Db 478 LRVLDVVLVWQDDEADGHMEREQILGLLVENLGHQATVRKGAALDCLRLVLAETAP- 534
Qy 279 HSRRTQYFYSWLLNVLLGLL-----VPVEDEHSTLLILGVLLTLRLYVLPLOOQVQKDT 331
Db 535 -----ETVMLGILDAGLTQKVPADSEHMGRLSCGVLLS-----VPALLQSIMHT 578
Qy 332 SLKSGFGVTRKEMEVSFAEOLVQVYELTLHHTQHDHNVVTGALELLQQLFRTPPPPELL 391
Db 579 A-----QRHRIVHSTVERV-----AHMDQVVOQEITVKVLSKIRELLGVHEFEI 624
Qy 392 -----OTLTAVGGIGOL--TAAKESGGRSGSIVELIAGG----- 427
Db 625 MGNVGRGDTLSRYQLSOVYGVSKSKTGGEVKAGARALPREQGWNGNNAVQNLDITLQ 684
Qy 428 --SSCSPVLSRKQKGVLLGEEEALED------SESRDVSSSALTAS 468
Db 685 IOSNCLDM-----GKIIMETEIKINDTVMRILEADTETEDGTTHQEPNIGAVIKVI 738
Qy 469 VKDEISGELAASSGVSTPGSA-----CHDII-----TEQPRS-----QH 502
Db 739 SDSELDPEHIKAGSVSEPTSGRLKRVTFGEIVKMTSPDASASTNNSRSTQTNQS 798
Qy 503 TLQADSVDLASCDLTSSATDGEEDILS-----HSSQVSAPVSDPAMD-NDGTOASS 555
Db 799 ITVSSSEDAALSDLVSS-----DEKSTLSRPVNDKRTTALVLEIFDNTKPLPQDRSLCSQ 854
Qy 556 PISDSOTTTEGDSAVTPSDSSEIVLDGTONQXGLQIGQPODEDEATGILPDEASEA 615
Db 855 PHRHQSNQASE-----PSSNS-----PRNRQENKARLSPSPSPG 889
Qy 616 PRNSS-----MALQOAHLLKNMSHCRQPSDSDSVKFLVRDEATEPGQENKPC 663
Db 890 FKRSTSTGTGNNNIPKVPKHQIEVLHNLQ--RDPSPRS-----QRSEDMGNDGKAL 942
Qy 664 RIKGDIGQSTDDSDAPLVHCVRLLSASFLLTAGKNVLVPRDRVRSVKA----- 712
Db 943 HATGNPSSS---PTQPLISRTSRASASTMSPVSPATPKSWEDLDI-VNLKTLLELRSGDWRN 998
Qy 713 -----LALSCVGAVALHP-----ESFFSKLYKVPPLDT 740
Db 999 RUMGIGOLELALSSSSNLAQVPLDLSLLRTLSSERHFEVSDKLRELLVNLISRLPLDN 1058
Qy 741 TEYPEQVYSDILNVIDHGDPOVGA---TAILCGTLICISILSRSRFH 785
Db 1059 LEERTPQILSGLCROGNAGANRVCKALMQRLPAGTIVAKLITSPEFLH 1105

```

```

RESULT 10
US-10-062-254-240
; Sequence 240, Application US/10062254
; Patent No. US20020138882A1
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Edgar B
; APPLICANT: Cahoon, Rebecca E
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Fang, Yiwon
; APPLICANT: Hantke, Sabine S.
; APPLICANT: Lee, Jfan-Ming
; APPLICANT: Li, Zhongsen
; APPLICANT: Miao, Guo-Hua
; APPLICANT: Morgante, Michele
; APPLICANT: Niu, Xiping

```



```

RESULT 14
US-09-946-805-4
; Sequence 4, Application US/09946805
; Patent No. US20020116734A1
; GENERAL INFORMATION:
; APPLICANT: Dickman, Martin B.
; TITLE OF INVENTION: PLANT DERIVED BAG HOMOLOGUES
; FILE REFERENCE: 480140.469
; CURRENT APPLICATION NUMBER: US/09/946,805
; CURRENT FILING DATE: 2001-09-04
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1043
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-946-805-4

Query Match          2.1%; Score 164; DB 10; Length 1043;
Best Local Similarity 20.3%; Pred. No. 0.042;

```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:14:29 : Search time 399,975 Seconds
(without alignments)
2487.216 Million cell updates/sec

Title: US-09-904-987-7

Perfect score: 7892

Sequence: 1 MATLEKLMKAFESLKSFQOQ.....DGIMASGRKASQPYPYELCSP 1543

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_Main:*

1: /cgn2_6/ptodata/1/paa/PTUS_COMB.pep.*

2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*

3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*

4: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*

5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*

6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*

7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*

8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*

9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*

10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*

11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*

12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*

13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*

14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*

15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*

16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*

17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*

18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*

19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*

20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*

21: /cgn2_6/ptodata/1/paa/US097_COMB.pep.*

22: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*

23: /cgn2_6/ptodata/1/paa/US099_COMB.pep.*

24: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*

25: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*

26: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*

27: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7892	100.0	1543	1	PCT-US02-22032-7
2	7892	100.0	1543	23	US-09-904-987-7
3	7833	99.3	3141	27	US-60-389-987-2250
4	7833	99.3	3141	27	US-60-412-418-2250
5	7821.5	99.1	3144	1	PCT-US99-05250-15
6	7821.5	99.1	3144	9	US-08-556-419A-1

7	7821.5	99.1	3144	9	US-08-556-419B-1	Sequence 1, Appli
8	7821.5	99.1	3144	16	US-09-212-271-21	Sequence 21, Appli
9	7821.5	99.1	3144	19	US-09-538-092-1118	Sequence 1118, Ap
10	7821.5	99.1	3223	1	PCT-US01-03800A-1777	Sequence 1777, Ap
11	7821.5	99.1	3223	1	PCT-US01-03800A-1840	Sequence 1840, Ap
12	7805.5	98.9	3144	4	US-08-085-000-6	Sequence 6, Appli
13	7057	89.4	3118	8	US-08-457-273A-8	Sequence 8, Appli
14	7057	89.4	3118	11	US-08-724-152-8	Sequence 31, Appli
15	2967.5	37.6	589	1	PCT-US99-05250-31	Sequence 30, Appli
16	2784.5	35.3	552	1	PCT-US99-05250-30	Sequence 29, Appli
17	2676.5	33.9	530	1	PCT-US99-05250-29	Sequence 28, Appli
18	2589.5	32.8	513	1	PCT-US99-05250-28	Sequence 3, Appli
19	884.5	11.2	171	1	PCT-US98-14376-3	Sequence 3, Appli
20	884.5	11.2	171	15	US-09-113-585-3	Sequence 21421, A
21	682	8.6	3528	27	US-60-167-217-21421	Sequence 426, App
22	679	8.6	3582	27	US-60-171-625-426	Sequence 17426, A
23	679	8.6	3582	27	US-60-173-464-17426	Sequence 21234, A
24	674.5	8.5	3583	20	US-09-614-150-21234	Sequence 21236, A
25	674.5	8.5	3583	27	US-60-191-637-21296	Sequence 16784, A
26	674.5	8.5	3583	27	US-60-191-681-16784	Sequence 2088, Ap
27	522	6.6	106	20	US-09-605-784A-2088	Sequence 6, Appli
28	488	6.2	94	18	US-09-463-874-6	Sequence 7, Appli
29	477.5	6.1	95	18	US-09-463-874-7	Sequence 8, Appli
30	477	6.0	96	18	US-09-463-874-8	Sequence 9, Appli
31	476.5	6.0	97	18	US-09-463-874-9	Sequence 10, Appli
32	476	6.0	98	18	US-09-463-874-10	Sequence 11, Appli
33	475.5	6.0	99	18	US-09-463-874-11	Sequence 12, Appli
34	475	6.0	100	18	US-09-463-874-12	Sequence 13, Appli
35	474.5	6.0	101	18	US-09-463-874-13	Sequence 14, Appli
36	474	6.0	102	18	US-09-463-874-14	Sequence 15, Appli
37	473.5	6.0	103	18	US-09-463-874-15	Sequence 16, Appli
38	473	6.0	104	18	US-09-463-874-16	Sequence 17, Appli
39	472.5	6.0	105	18	US-09-463-874-17	Sequence 18, Appli
40	472	6.0	106	18	US-09-463-874-18	Sequence 19, Appli
41	471.5	6.0	107	18	US-09-463-874-19	Sequence 20, Appli
42	471	6.0	108	18	US-09-463-874-20	Sequence 21, Appli
43	470.5	6.0	109	18	US-09-463-874-21	Sequence 22, Appli
44	470	6.0	110	18	US-09-463-874-22	Sequence 23, Appli
45	469.5	5.9	111	18	US-09-463-874-23	

ALIGNMENTS

RESULT 1

PCT-US02-22032-7

; Sequence 7, Application PC/TUS0222032

; GENERAL INFORMATION:

; APPLICANT: Novactyl, Inc.

; TITLE OF INVENTION: Methods and Compositions for Controlling Protein Assembly or A

; FILE REFERENCE: 42108/34520

; CURRENT APPLICATION NUMBER: PCT/US02/22032

; CURRENT FILING DATE: 2002-07-11

; PRIOR APPLICATION NUMBER: 09/904,987

; PRIOR FILING DATE: 12 July 2001

; NUMBER OF SEQ ID NOS: 7

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 7

; LENGTH: 1543

; TYPE: PRT

; ORGANISM: homo sapiens

; PUBLICATION INFORMATION:

; DATABASE ACCESSION NUMBER: NCBI ENTREZ / XP_003405

; DATABASE ENTRY DATE: 2001-04-16

; RELEVANT RESIDUES: (1)..(1543)

PCT-US02-22032-7

Query Match 100.0%; Score 7892; DB 1; Length 1543;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1543; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATLEKLMKAFESLKSFQOQ.....DGIMASGRKASQPYPYPPLPQPPQAQPL 60

|||||

Db 241 FANDNEIKVLLKAFITANLKSSPTIRRTAAGSAVSICQHSRRTQYFYSWLLNVLLGLVP 300
QY 301 VEDEHSTLLILGVLTLRYLPLVLPLOQQVKDTSKGSFQVTRKEMEVSASQLVQVYELT 360
Db 301 VEDEHSTLLILGVLTLRYLPLVLPLOQQVKDTSKGSFQVTRKEMEVSASQLVQVYELT 360
QY 361 LHHTQHODHNVVTGALELLQQLFRTPPPELLQTLTAVGGIGOLTAAKEESGGRSGSIV 420
Db 361 LHHTQHODHNVVTGALELLQQLFRTPPPELLQTLTAVGGIGOLTAAKEESGGRSGSIV 420
QY 421 ELIAGGSGSCSPVLSRKQKGVLLGEEBEEALDDSESRSVSSSALTASVKDEISGELAAS 480
Db 421 ELIAGGSGSCSPVLSRKQKGVLLGEEBEEALDDSESRSVSSSALTASVKDEISGELAAS 480
QY 481 SGVSTPGSAGHDITEQPRSOHTLQADSVDLASCDLTSSATDGDDEEDILSHSSQVSQAVP 540
Db 481 SGVSTPGSAGHDITEQPRSOHTLQADSVDLASCDLTSSATDGDDEEDILSHSSQVSQAVP 540
QY 541 SDPAMDNDGTQASSPISDSSQTTTEGPDASVTPSDSEIVLDTNOYLGLOIGQPODE 600
Db 541 SDPAMDNDGTQASSPISDSSQTTTEGPDASVTPSDSEIVLDTNOYLGLOIGQPODE 600
QY 601 DEEATGILPDASAFRNSMALQQAHLKMNHSHCRQPSDSSVDKFLVRDEATEPGQDEN 660
Db 601 DEEATGILPDASAFRNSMALQQAHLKMNHSHCRQPSDSSVDKFLVRDEATEPGQDEN 660
QY 661 KPCRKIGDGGSTDDSDAPLVHCVRLLSASFLLTGKGNLVLPDRDVRVSKALALSCVGA 720
Db 661 KPCRKIGDGGSTDDSDAPLVHCVRLLSASFLLTGKGNLVLPDRDVRVSKALALSCVGA 720
QY 721 AVALHPESFFSKLVPLDTTEYPEEQVSDILNYIDHGDPPQVRGATAILCGTLICLSILS 780
Db 721 AVALHPESFFSKLVPLDTTEYPEEQVSDILNYIDHGDPPQVRGATAILCGTLICLSILS 780
QY 781 RSRFHVGWGMCTIRLTGNTFSLADCIPLLRKTLKDESSVTCKLACTAVRNCVMSLCSSS 840
Db 781 RSRFHVGWGMCTIRLTGNTFSLADCIPLLRKTLKDESSVTCKLACTAVRNCVMSLCSSS 840
QY 841 YSELGLQLIIDVLTRNSSVWLVRTELETLAEIDFRLVSFLEAKAENLHRGAHHTYGLL 900
Db 841 YSELGLQLIIDVLTRNSSVWLVRTELETLAEIDFRLVSFLEAKAENLHRGAHHTYGLL 900
QY 901 KLOERVLNNAVYHLLGDEDPVRHVAASLIRLVKLFYKCDGQADPPVAVARDQSSVY 960
Db 901 KLOERVLNNAVYHLLGDEDPVRHVAASLIRLVKLFYKCDGQADPPVAVARDQSSVY 960
QY 961 LKLLMHETOPPSHFSVSTITRIYRGYNLPSITDVTMNNLSRVIAAVSHELITSTTRAL 1020
Db 961 LKLLMHETOPPSHFSVSTITRIYRGYNLPSITDVTMNNLSRVIAAVSHELITSTTRAL 1020
QY 1021 TFGCEALCLLSTAFFPVCINSLGWHGCVPLPSASDESRSKCTVGMATMILTLLSSAWFPL 1080
Db 1021 TFGCEALCLLSTAFFPVCINSLGWHGCVPLPSASDESRSKCTVGMATMILTLLSSAWFPL 1080
QY 1081 DLSAHODALLIAGNLLAASAPKSLRWSAEEBANPAATKQEEWPALGDALVPVMEQL 1140
Db 1081 DLSAHODALLIAGNLLAASAPKSLRWSAEEBANPAATKQEEWPALGDALVPVMEQL 1140
QY 1141 FSHLLKVINICAHVLDVAPGPAIKALPSLTNPBSLSPIRRRKKEPEGEQASVPLSPK 1200
Db 1141 FSHLLKVINICAHVLDVAPGPAIKALPSLTNPBSLSPIRRRKKEPEGEQASVPLSPK 1200
QY 1201 KGSEASARSQSDTSGPVTTSKSSLSGFSYHLPSYLLKHDVLTATHYKVTLDLONSTE 1260
Db 1201 KGSEASARSQSDTSGPVTTSKSSLSGFSYHLPSYLLKHDVLTATHYKVTLDLONSTE 1260
QY 1261 KFGFLRSALDVLQSILELATLQIDKCVBEELIGYUKSCFSRPPMMATVCVOQLLKTFLG 1320
Db 1261 KFGFLRSALDVLQSILELATLQIDKCVBEELIGYUKSCFSRPPMMATVCVOQLLKTFLG 1320
QY 1321 TNLASQFDGLSSNPKSQGRRAQLGSSSVRPGLYHYCFMAPYTHFTQALADASLRNMVQA 1380

Db 1321 TNLASQFDGLSSNPKSQGRRAQLGSSSVRPGLYHYCFMAPYTHFTQALADASLRNMVQA 1380
QY 1381 EQENDTSGWFDVLQKVSTQLKTNLTSTVKNRADKNATHNHIRLFEPLVLIKALKQYTTTTT 1440
Db 1381 EQENDTSGWFDVLQKVSTQLKTNLTSTVKNRADKNATHNHIRLFEPLVLIKALKQYTTTTT 1440
QY 1441 VOLOKQVLDLLAQLVQLRVNYCCLLSDQVFIGVLFKQFIEVQGFRESAIIPIFFFL 1500
Db 1441 VOLOKQVLDLLAQLVQLRVNYCCLLSDQVFIGVLFKQFIEVQGFRESAIIPIFFFL 1500
QY 1501 VLLSYERYHSHQIIGIPKIIQLCDGIMASGRKASPOPYRLCSP 1543
Db 1501 VLLSYERYHSHQIIGIPKIIQLCDGIMASGRKASPOPYRLCSP 1543

RESULT 3

US-60-389-987-2250
; Sequence 2250, Application US/60389987
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; CURRENT APPLICATION NUMBER: US/60/389, 987
; CURRENT FILING DATE: 2002-06-17
; NUMBER OF SEQ ID NOS: 3025
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2250
; LENGTH: 3141
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-389-987-2250

Query Match 99.3%; Score 7833; DB 27; Length 3141;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1533; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATLEKLMKAFESLKSFOQQQQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPPPPPPPPP 60
Db 1 MATLEKLMKAFESLKSFOQQQQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPPPPPPPPP 60
QY 61 LPQPQPP 120
Db 61 LPQPQPP 120
QY 121 QKLLGTAMBLFLCSDDAESDVRMVADECLNKVIALMDNSLPRLOLEYLKEIKKNGAPR 180
Db 121 QKLLGTAMBLFLCSDDAESDVRMVADECLNKVIALMDNSLPRLOLEYLKEIKKNGAPR 180
QY 181 SURAAWLREAEHLAVRPQKCRPYLVNLLPCLTRTSKPEESVOETLAAAVPKIMASFGN 240
Db 181 SURAAWLREAEHLAVRPQKCRPYLVNLLPCLTRTSKPEESVOETLAAAVPKIMASFGN 240
QY 241 FANDNEIKVLLKAFITANLKSSSPTIRRTAAGSAVSICQHSRRTQYFYSWLLNVLLGLVP 300
Db 241 FANDNEIKVLLKAFITANLKSSSPTIRRTAAGSAVSICQHSRRTQYFYSWLLNVLLGLVP 300
QY 301 VEDEHSTLLILGVLTLRYLPLVLPLOQQVKDTSKGSFQVTRKEMEVSASQLVQVYELT 360
Db 301 VEDEHSTLLILGVLTLRYLPLVLPLOQQVKDTSKGSFQVTRKEMEVSASQLVQVYELT 360
QY 361 LHHTQHODHNVVTGALELLQQLFRTPPPELLQTLTAVGGIGOLTAAKEESGGRSGSIV 420
Db 361 LHHTQHODHNVVTGALELLQQLFRTPPPELLQTLTAVGGIGOLTAAKEESGGRSGSIV 420
QY 421 ELIAGGSGSCSPVLSRKQKGVLLGEEBEEALDDSESRSVSSSALTASVKDEISGELAAS 480

Db 541 AVPSDPAMDINDGTQASSPISDSSQTTEGPDQSAVTPSDSSSEIVLDGTDNQYLGQIGQP 600
Qy 598 QDEDEATGILPDEASEAFRNSMALQQAHLKKNSHCRQPSDSSVDKFLVLRDEATEPGD 657
Db 601 QDEDEATGILPDEASEAFRNSMALQQAHLKKNSHCRQPSDSSVDKFLVLRDEATEPGD 660
Qy 658 QENKPCRIGDQIGQSDDDSDAPLVHCVRLLSASFLLTGKKNVLPDRDVRVSVKALALSC 717
Db 661 QENKPCRIGDQIGQSDDDSDAPLVHCVRLLSASFLLTGKKNVLPDRDVRVSVKALALSC 720
Qy 718 VGAVALHPESFFSKLYKVPDTPTEPEQYVSDILNYIDHGDQVPGATAILCGTLICS 777
Db 721 VGAVALHPESFFSKLYKVPDTPTEPEQYVSDILNYIDHGDQVPGATAILCGTLICS 780
Qy 778 ILSRFRHVGDMGTTTTLTGNTFSLADCIPLLRKTLKDESSVTCCLACTAVRNCVMSLC 837
Db 781 ILSRFRHVGDMGTTTTLTGNTFSLADCIPLLRKTLKDESSVTCCLACTAVRNCVMSLC 840
Qy 838 SSSYSELGLQLIIDVLTNRSSVWLVRTLETLAEIDFRLVSFLKAEKAEHLRGAHHYT 897
Db 841 SSSYSELGLQLIIDVLTNRSSVWLVRTLETLAEIDFRLVSFLKAEKAEHLRGAHHYT 900
Qy 898 GLLKQERVLNNVVIHLLGDEDPVRHVAASLIRLVPLKFKYKCDQGGQADPVVAVARDQS 957
Db 901 GLLKQERVLNNVVIHLLGDEDPVRHVAASLIRLVPLKFKYKCDQGGQADPVVAVARDQS 960
Qy 958 SVYLLKLMHETQPSHFSYSTIRIYRGYNLLPSITDVTMNNLSRVIAAVSHELITSTT 1017
Db 961 SVYLLKLMHETQPSHFSYSTIRIYRGYNLLPSITDVTMNNLSRVIAAVSHELITSTT 1020
Qy 1018 RALTFCCEALCLSTAFFPVCIVSLGWCHGVPPLSDASDSRKSCTVGMATMILTLLSSAW 1077
Db 1021 RALTFCCEALCLSTAFFPVCIVSLGWCHGVPPLSDASDSRKSCTVGMATMILTLLSSAW 1080
Qy 1078 FPLDLAHOALILAGNLIAASAPKSLRWSWASEEENPAATKQEEVWPALGDRLVPMV 1137
Db 1081 FPLDLAHOALILAGNLIAASAPKSLRWSWASEEENPAATKQEEVWPALGDRLVPMV 1140
Qy 1138 EQLFSLHLLKVINICAHVLDVAPGAIPAALPSLTNPSSLSPIRRKGEKEPEQASVPL 1197
Db 1141 EQLFSLHLLKVINICAHVLDVAPGAIPAALPSLTNPSSLSPIRRKGEKEPEQASVPL 1200
Qy 1198 SPKKGSEASARSQDSGTSGVTTSKSSSLGSHFYHPLSYLKLHDVLKATHANKVTLIDLQN 1257
Db 1201 SPKKGSEASARSQDSGTSGVTTSKSSSLGSHFYHPLSYLKLHDVLKATHANKVTLIDLQN 1260
Qy 1258 STEKFGFLRSALDVLISOILELATLQDICKVBEIILGYLKSCFSREPMMATVCVQQLKT 1317
Db 1261 STEKFGFLRSALDVLISOILELATLQDICKVBEIILGYLKSCFSREPMMATVCVQQLKT 1320
Qy 1318 LFGTNLASQFDGLSSNPKSQGQRAQLGSSSVRPGLYHYCFMAYPTHFTQALADASLRNM 1377
Db 1321 LFGTNLASQFDGLSSNPKSQGQRAQLGSSSVRPGLYHYCFMAYPTHFTQALADASLRNM 1380
Qy 1378 VQAEQENDTSGWFDVLQKYSTQTLNLTSTVTKNRADKNAIHNRILFEPLVLIKALKQYTT 1437
Db 1381 VQAEQENDTSGWFDVLQKYSTQTLNLTSTVTKNRADKNAIHNRILFEPLVLIKALKQYTT 1440
Qy 1438 TTCVQLQKQVLDLLAQLVQLRVNVCLLDSQVFIGFVLQKFIEYVQGFRESEAIIPNIF 1497
Db 1441 TTCVQLQKQVLDLLAQLVQLRVNVCLLDSQVFIGFVLQKFIEYVQGFRESEAIIPNIF 1500
Qy 1498 FFLVLLSYERHSKQIIGIPKIIQLCDGIMASGRKA 1533
Db 1501 FFLVLLSYERHSKQIIGIPKIIQLCDGIMASGRKA 1536

RESULT 6

US-08-556-419A-1

; Sequence 1: Application us/08556419A

; GENERAL INFORMATION:

; APPLICANT: Ross, Christopher A

; APPLICANT: Li, Xiao-Jiang

; APPLICANT: Li, Shi-Hua
; APPLICANT: Sharp, Alan H
; APPLICANT: Lanahan, Anthony
; APPLICANT: Worley, Paul F
; APPLICANT: Snyder, Solomon H
; TITLE OF INVENTION: Huntingtin-Associated Protein
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Allegretti, LTD.
; STREET: 1001 G Street, NW suite 1100
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/556,419A
; FILING DATE:
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Kagan, Sarah A
; REGISTRATION NUMBER: 32141
; REFERENCE/DOCKET NUMBER: 1107.52271
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202.508.9100
; TELEFAX: 202.508.9299
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3144 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; US-08-556-419A-1

Query Match 99.18; Score 7821.5; DB 9; Length 3144;

Best Local Similarity 99.88; Pred. No. 0;

Matches 1533; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

Qy 1 MATLEKLMKAFESLKSF---OQQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPQA 57
Db 1 MATLEKLMKAFESLKSFOQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPQA 60
Qy 58 QPLLPQPP 117
Db 61 QPLLPQPP 120
Qy 118 PEFQKLLGIAMFELFLCSDDAESDVRWVADECLANKVIKALMDSNLPRLQLEYKEIKNG 177
Db 121 PEFQKLLGIAMFELFLCSDDAESDVRWVADECLANKVIKALMDSNLPRLQLEYKEIKNG 180
Qy 178 APRSLRAALWFAELAHVLPQKCRPYLVNLLPCLITRTSKRPEESVQETLAAAVPKIMAS 237
Db 181 APRSLRAALWFAELAHVLPQKCRPYLVNLLPCLITRTSKRPEESVQETLAAAVPKIMAS 240
Qy 238 FGNFANDNEIKVLLKAFATANKLSSPTTRTAAGSAVSVICQHSRRTQYFYSWLLNVLLGL 297
Db 241 FGNFANDNEIKVLLKAFATANKLSSPTTRTAAGSAVSVICQHSRRTQYFYSWLLNVLLGL 300
Qy 298 LVPVEDEHSTLLILGLVLTLYLVLPLOQQVKTSLKSGFVTRKEMVSPSABQLVOVY 357
Db 301 LVPVEDEHSTLLILGLVLTLYLVLPLOQQVKTSLKSGFVTRKEMVSPSABQLVOVY 360
Qy 358 ELTLHHTQHQHNVTGALELILQQLFRTPPPPELLQTLTAVGGIGQLTAAKESGGRSRG 417
Db 361 ELTLHHTQHQHNVTGALELILQQLFRTPPPPELLQTLTAVGGIGQLTAAKESGGRSRG 420
Qy 418 SIVELIAGGGSCSPVLSRKQKQKGVLLGEEALEDDDSERSDVSASSALTASVKDEISEL 477

Db 421 STVELIAGGSCSPVLSRKQKGVLLGEEALEDDSESVDVSSALTSVKDISEL 480
QY 478 AASSGVSTPGSAGHDIITPEQRSQHTLQADSVDLASCOLDTSSATGDDEEDILSHSSQVS 537
Db 481 AASSGVSTPGSAGHDIITPEQRSQHTLQADSVDLASCOLDTSSATGDDEEDILSHSSQVS 540
QY 538 AVPSOPAMDLNDGTQASSPIDSSOTTTEGPDASVTPSDSSSEIVLDGTDNOYGLGQICQP 597
Db 541 AVPSOPAMDLNDGTQASSPIDSSOTTTEGPDASVTPSDSSSEIVLDGTDNOYGLGQICQP 600
QY 598 QDEDEATGILPDEASEAFRNSSMALQQAHLKNNSHCRQPSDSSVDKFLVRDEATEPGD 657
Db 601 QDEDEATGILPDEASEAFRNSSMALQQAHLKNNSHCRQPSDSSVDKFLVRDEATEPGD 660
QY 658 QENKPCRIGDITGQSTDDSDAPLVHCVRLLNSASFLTTGGKNVLPDRVRVSVKALALSC 717
Db 661 QENKPCRIGDITGQSTDDSDAPLVHCVRLLNSASFLTTGGKNVLPDRVRVSVKALALSC 720
QY 718 VGAVALHPESFESKLYKVPDITTEYPEEQVSDILNVIDHGDPQVGCATAILCGTLICS 777
Db 721 VGAVALHPESFESKLYKVPDITTEYPEEQVSDILNVIDHGDPQVGCATAILCGTLICS 780
QY 778 ILSRFRHVGMGTIRLTGNTFTSLADCIPLLRKTLDKDESSVTKLACTAVRNCVMSLC 837
Db 781 ILSRFRHVGMGTIRLTGNTFTSLADCIPLLRKTLDKDESSVTKLACTAVRNCVMSLC 840
QY 838 SSSYSELGQLIIDLVTLRNSSYWLVRTELLETLAEIDFRVLVSLEAKAENLHRAHHT 897
Db 841 SSSYSELGQLIIDLVTLRNSSYWLVRTELLETLAEIDFRVLVSLEAKAENLHRAHHT 900
QY 898 GLKQLQERVNNVTHLLGDEDPVRVHVAASLIRLVPKFKYKDOGOADPVAVARDQS 957
Db 901 GLKQLQERVNNVTHLLGDEDPVRVHVAASLIRLVPKFKYKDOGOADPVAVARDQS 960
QY 958 SVYLLMHETQPPSHFVSITRIYRGNLLPSTDTVTMENNLSRVIAAASHELITST 1017
Db 961 SVYLLMHETQPPSHFVSITRIYRGNLLPSTDTVTMENNLSRVIAAASHELITST 1020
QY 1018 RALTFGCCALCLLSTAPVCVWSLGHGCVPPISASDESRSKCTVGMATMILLLSAW 1077
Db 1021 RALTFGCCALCLLSTAPVCVWSLGHGCVPPISASDESRSKCTVGMATMILLLSAW 1080
QY 1078 PDLDSAHODALILAGNLLAASAPKSLRSSWASEEENAPAAKOEYVMPALGDRLVPMV 1137
Db 1081 PDLDSAHODALILAGNLLAASAPKSLRSSWASEEENAPAAKOEYVMPALGDRLVPMV 1140
QY 1138 EQLFSLHLKVINICAHVLDVAPGPAIKAAPLSLTNPSPSLPIRRKKEKEPGEQASVPL 1197
Db 1141 EQLFSLHLKVINICAHVLDVAPGPAIKAAPLSLTNPSPSLPIRRKKEKEPGEQASVPL 1200
QY 1198 SPKKGSEASASROSSTSGPVTTSKSSSLGSGFYHLPSYKLHDLVKATHANYKVTLDLQ 1257
Db 1201 SPKKGSEASASROSSTSGPVTTSKSSSLGSGFYHLPSYKLHDLVKATHANYKVTLDLQ 1260
QY 1258 STEKFGGLRSALDVLQILELATLQDIDGKCVBEILGYLKSCEPMPMATVCVQQLLKT 1317
Db 1261 STEKFGGLRSALDVLQILELATLQDIDGKCVBEILGYLKSCEPMPMATVCVQQLLKT 1320
QY 1318 LFGTNLASQFGLSNPSKSGRAQRGLGSSVRPGLHYCPMAYTHFTQALADASLRNM 1377
Db 1321 LFGTNLASQFGLSNPSKSGRAQRGLGSSVRPGLHYCPMAYTHFTQALADASLRNM 1380
QY 1378 VQAEQENDTSGWFDVQKVSQTLKNTLTSVTKNRADKNAIHNHRLRLEPLVKAQKQYTT 1437
Db 1381 VQAEQENDTSGWFDVQKVSQTLKNTLTSVTKNRADKNAIHNHRLRLEPLVKAQKQYTT 1440
QY 1438 TTCVOLQKQVLDLQAQLVQLRVNYCLLDSQVFIQKFEYIEVQGFRESEAIIPNIF 1497
Db 1441 TTCVOLQKQVLDLQAQLVQLRVNYCLLDSQVFIQKFEYIEVQGFRESEAIIPNIF 1500
QY 1498 FFLVLLSYERYHSKQIIGIPKIIQICDGMASGRKA 1533
|||||

Db 1501 FFLVLLSYERYHSKQIIGIPKIIQICDGMASGRKA 1536

RESULT 7

US-08-556-419B-1
: Sequence 1, Application US/08556419B
: GENERAL INFORMATION:
: APPLICANT: Ross, Christopher A
: APPLICANT: Li, Xiao-Jiang
: APPLICANT: Li, Shi-Hua
: APPLICANT: Sharp, Alan H
: APPLICANT: Lanahan, Anthony
: APPLICANT: Worley, Paul F
: APPLICANT: Snyder, Solomon H
: TITLE OF INVENTION: Huntingtin-Associated Protein
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Banner & Witcoff, LTD.
: STREET: 1001 G Street, NW suite 1100
: CITY: Washington
: STATE: DC
: COUNTRY: USA
: ZIP: 20001
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/556,419B
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Kagan, Sarah A
: REGISTRATION NUMBER: 32141
: REFERENCE/DOCKET NUMBER: 1107.52271
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 202.508.9100
: TELEFAX: 202.508.9299
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3144 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: Protein
: ORIGINAL SOURCE:
: ORGANISM: Homo sapiens
: US-08-556-419B-1

Query Match 99.1%; Score 7821.5; DB 9; Length 3144;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1533; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 1 MATLEKLMKAFESLKSF--QQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPQPPQA 57
Db 1 MATLEKLMKAFESLKSFQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPQPPQA 60
QY 58 QPLLPOQPP 117
Db 61 QPLLPOQPP 120
QY 118 PEFQKLLGIAAMELFLCSDDAESDVRMVADECLNKVIKALMDSNLPRLQLELYKEIKNG 177
Db 121 PEFQKLLGIAAMELFLCSDDAESDVRMVADECLNKVIKALMDSNLPRLQLELYKEIKNG 180
QY 178 APRSLRAALWRFRAELAHVLRPQRCRPLYNNLLPCLTRTSKRPEESVQETLAAAVPKIMAS 237
Db 181 APRSLRAALWRFRAELAHVLRPQRCRPLYNNLLPCLTRTSKRPEESVQETLAAAVPKIMAS 240
QY 238 FGNFANDNEIKVLLKAFIANLKSSSPTIRRTAAGSAVSCQHSRRRTQYFYSNLLNVLLGL 297
Db 241 FGNFANDNEIKVLLKAFIANLKSSSPTIRRTAAGSAVSCQHSRRRTQYFYSNLLNVLLGL 300

Qy 298 LVPVEDEHSTLLILGLVLTLLRYLVLPLQOQVKDTSLKSGFGVTRKMEVSPASQOLVQVY 357
Db 301 LVPVEDEHSTLLILGLVLTLLRYLVLPLQOQVKDTSLKSGFGVTRKMEVSPASQOLVQVY 360
Qy 358 ELTLHHTQHODHNVTGALQLLOQLFRTPPPPELLQTLTAVGGICOLTAAKEESGRSRG 417
Db 361 ELTLHHTQHODHNVTGALQLLOQLFRTPPPPELLQTLTAVGGICOLTAAKEESGRSRG 420
Qy 418 SIVELIAGGSSCPVLSRKQKQKGVLLGEEALEDDSESRSVSSSALTASVKDEISEL 477
Db 421 SIVELIAGGSSCPVLSRKQKQKGVLLGEEALEDDSESRSVSSSALTASVKDEISEL 480
Qy 478 AASSGVSTPGSAGHDITTEQPRSQHTLQADSVDLASCDLTSSATDGDDEEDILSHSSQVS 537
Db 481 AASSGVSTPGSAGHDITTEQPRSQHTLQADSVDLASCDLTSSATDGDDEEDILSHSSQVS 540
Qy 538 AVPSOPAMDLNQTQASSPIDSSQTTTEGPDASVTPSDSSEIVLDGTDNOYLGILQIQP 597
Db 541 AVPSOPAMDLNQTQASSPIDSSQTTTEGPDASVTPSDSSEIVLDGTDNOYLGILQIQP 600
Qy 598 QDEDEATGILPDEASEAFRNSMALQQAHLKNNKSHCRQPSDSSVDKFLVRDEATEPGD 657
Db 601 QDEDEATGILPDEASEAFRNSMALQQAHLKNNKSHCRQPSDSSVDKFLVRDEATEPGD 660
Qy 658 QENKPCRIGDQSTDDSDAPLVHCVRLLSASFLTGTGKNVLPDRDVRVSVKALASC 717
Db 661 QENKPCRIGDQSTDDSDAPLVHCVRLLSASFLTGTGKNVLPDRDVRVSVKALASC 720
Qy 718 VGAVALHPESFESKLYKVPDLTTEPEQYVSDILNVIDHGDPQVGTALILGTLICS 777
Db 721 VGAVALHPESFESKLYKVPDLTTEPEQYVSDILNVIDHGDPQVGTALILGTLICS 780
Qy 778 ILSRSRHFVGDWNGTIRTLTGNTFSLADICPLRLKTLKDESSVTCKLACTAVRNCVMSLC 837
Db 781 ILSRSRHFVGDWNGTIRTLTGNTFSLADICPLRLKTLKDESSVTCKLACTAVRNCVMSLC 840
Qy 838 SSSYSELGLQILIDVTLRNSYWLVRTELLETLAEIDFRLVSFLAEAKENLHRCAGHYT 897
Db 841 SSSYSELGLQILIDVTLRNSYWLVRTELLETLAEIDFRLVSFLAEAKENLHRCAGHYT 900
Qy 898 GLLKQERVLNNVTHLLGDEDPVRHVAASLIRLPKLFYKCDQGOADPVAVARQDS 957
Db 901 GLLKQERVLNNVTHLLGDEDPVRHVAASLIRLPKLFYKCDQGOADPVAVARQDS 960
Qy 958 SVYKLMLHETOPPSHFSYSTIRYGNLPLSITDVTMNNLSRVIAAASVHELITSTT 1017
Db 961 SVYKLMLHETOPPSHFSYSTIRYGNLPLSITDVTMNNLSRVIAAASVHELITSTT 1020
Qy 1018 RALTFCCEALCLLSTAFVPCVWSLGHGCVPLPSASDESRSKCTVGMATMILLSSAW 1077
Db 1021 RALTFCCEALCLLSTAFVPCVWSLGHGCVPLPSASDESRSKCTVGMATMILLSSAW 1080
Qy 1078 FPLDLSAHQDALLAGNLLAASAPKSLRSSWASEEAPATKQEEVWPALGDRLVPMV 1137
Db 1081 FPLDLSAHQDALLAGNLLAASAPKSLRSSWASEEAPATKQEEVWPALGDRLVPMV 1140
Qy 1138 EQLFSLHLLKVINICAHVLDVDPGPAIKAAPLSLTNPPLSIPRRKKEKEPEQASVPL 1197
Db 1141 EQLFSLHLLKVINICAHVLDVDPGPAIKAAPLSLTNPPLSIPRRKKEKEPEQASVPL 1200
Qy 1198 SPKKGSEASAAQSQSTGPTVTSKSSSLGSGFYHLPSYLKLDHVLKATHANYKVTLIDLQ 1257
Db 1201 SPKKGSEASAAQSQSTGPTVTSKSSSLGSGFYHLPSYLKLDHVLKATHANYKVTLIDLQ 1260
Qy 1258 STEKFGFLRSALDVLSQILELATLQDICKVVEILGYLKSCFSREPMATVCVQOQLKT 1317
Db 1261 STEKFGFLRSALDVLSQILELATLQDICKVVEILGYLKSCFSREPMATVCVQOQLKT 1320
Qy 1318 LFGTNLASQFDGLSSNPSKQRAQRLGSSSVRPGLYHYCFMAYTHFTQALADASLRNM 1377
Db 1321 LFGTNLASQFDGLSSNPSKQRAQRLGSSSVRPGLYHYCFMAYTHFTQALADASLRNM 1380
Qy 1378 VQAEQENDTSGWFDVLQKYSTQLKNTLTSVTKNRADKNAIHNHIRLFEPLVTKALKQYTT 1437

Db 1381 VQAEQENDTSGWFDVLQKYSTQLKNTLTSVTKNRADKNAIHNHIRLFEPLVTKALKQYTT 1440
Qy 1438 TTCVQLOKQVLDLLAQLVQLRVNYCLLDSQVFTGFVLFKQFIEVQGFRESAIIIPNIF 1497
Db 1441 TTCVQLOKQVLDLLAQLVQLRVNYCLLDSQVFTGFVLFKQFIEVQGFRESAIIIPNIF 1500
Qy 1498 FFLVLLSYERYHSHKQIIGIPKIIQLCDGIMASGRKA 1533
Db 1501 FFLVLLSYERYHSHKQIIGIPKIIQLCDGIMASGRKA 1536

RESULT 8
US-09-212-271-21
; Sequence 21, Application US/09212271
; GENERAL INFORMATION:
; APPLICANT: Ross, Christopher
; APPLICANT: Li, Xiao-Jiang
; APPLICANT: Li, Shi-Hua
; APPLICANT: Sharp, Alan
; APPLICANT: Lanahan, Anthony
; APPLICANT: Worley, Paul
; APPLICANT: Snyder, Solomon
; TITLE OF INVENTION: Huntingtin-associated protein
; FILE REFERENCE: 01107.77898
; CURRENT APPLICATION NUMBER: US/09/212, 271
; CURRENT FILING DATE: 1998-12-16
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21
; LENGTH: 3144
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-212-271-21

Query Match 99.1%; Score 7821.5; DB 16; Length 3144;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1533; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

Qy 1 MATLEKLMKAFESLKSF---QQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPQ 57
Db 1 MATLEKLMKAFESLKSFQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPQ 60
Qy 58 QLLPQQPP 117
Db 61 QLLPQQPP 120
Qy 118 PEFQKLLGIAMFELFLCSDDAESDVRMVADCLNKVTKALMDSNLPRLQLELYKEIKNG 177
Db 121 PEFQKLLGIAMFELFLCSDDAESDVRMVADCLNKVTKALMDSNLPRLQLELYKEIKNG 180
Qy 178 APRSLRAALWRFABLAHLVRPQKCRPYLVNLLPCLTRTSKRPEESVQETLAAAVPKIMAS 237
Db 181 APRSLRAALWRFABLAHLVRPQKCRPYLVNLLPCLTRTSKRPEESVQETLAAAVPKIMAS 240
Qy 238 FGNFANDNEIKVLLKAFIANKLKSSPTIRRTAAGSAVSIQHSRRTQYFYSWLLNVLGL 297
Db 241 FGNFANDNEIKVLLKAFIANKLKSSPTIRRTAAGSAVSIQHSRRTQYFYSWLLNVLGL 300
Qy 298 LVPVEDEHSTLLILGLVLTLLRYLVLPLQOQVKDTSLKSGFGVTRKMEVSPASQOLVQVY 357
Db 301 LVPVEDEHSTLLILGLVLTLLRYLVLPLQOQVKDTSLKSGFGVTRKMEVSPASQOLVQVY 360
Qy 358 ELTLHHTQHODHNVTGALQLLOQLFRTPPPPELLQTLTAVGGICOLTAAKEESGRSRG 417
Db 361 ELTLHHTQHODHNVTGALQLLOQLFRTPPPPELLQTLTAVGGICOLTAAKEESGRSRG 420
Qy 418 SIVELIAGGSSCPVLSRKQKQKGVLLGEEALEDDSESRSVSSSALTASVKDEISEL 477
Db 421 SIVELIAGGSSCPVLSRKQKQKGVLLGEEALEDDSESRSVSSSALTASVKDEISEL 480
Qy 478 AASSGVSTPGSAGHDITTEQPRSQHTLQADSVDLASCDLTSSATDGDDEEDILSHSSQVS 537

[illegible]

RESULT 14

```

1  ; Sequence 8, Application US/08724152
2  ; GENERAL INFORMATION:
3  ; APPLICANT: Hayden, Michael
4  ; APPLICANT: Lin, Biaoyang
5  ; APPLICANT: Nasir, Jamal
6  ; TITLE OF INVENTION: MOUSE MODEL FOR HUNTINGTON'S
7  ; TITLE OF INVENTION: DISEASE AND
8  ; TITLE OF INVENTION: RELATED DNA SEQUENCES
9  ; NUMBER OF SEQUENCES: 8
10 ; CORRESPONDENCE ADDRESS:
11 ; ADDRESSEE: Bell, Seltzer, Park & Gibson
12 ; STREET: 1211 East Morehead Street,
13 ; CITY: Charlotte,
14 ; STATE: North Carolina
15 ; COUNTRY: U.S.A.
16 ; ZIP: 34009
17 ; COMPUTER READABLE FORM:
18 ; MEDIUM TYPE: Floppy disk
19 ; COMPUTER: IBM PC compatible
20 ; OPERATING SYSTEM: PC-DOS/MS-DOS
21 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
22 ; CURRENT APPLICATION DATA:
23 ; APPLICATION NUMBER: US/08/724,152
24 ; FILING DATE:
25 ; CLASSIFICATION: 800
26 ; PRIOR APPLICATION DATA:
27 ; APPLICATION NUMBER: US/08/127,971
28 ; FILING DATE:
29 ; ATTORNEY/AGENT INFORMATION:
30 ; NAME: Layton, Jr., Samuel G
31 ; REGISTRATION NUMBER: 22807
32 ; REFERENCE/DOCKET NUMBER: 3477-85
33 ; TELECOMMUNICATION INFORMATION:
34 ; TELEPHONE: 704-331-6000
35 ; TELEFAX: 704-334-2014

```

```
QY 780 SRSRHFVGDWMTGIRTGNTESLADCIPLLRKTLKDESSVTCKLACTAVRNCVMSLCS 839
Db 760 SRSRLRGVDMGLNIRTLGNTESLVDICPLLRKTLKDESSVTCKLACTAVRNCVMSLCS 819
QY 840 SYSELGQLIIDIIVLRNSYWLVRTELLETLAEIDFRVLSPLEAKAENLHRGAHHYTG 899
Db 820 SYSDLGQLIDMLPLKNSYWLVRTELLETLAEIDFRVLSPLEAKAENLHRGAHHYTG 879
QY 900 LKQERVLNNVYIHLGDEDPVRHVAASLRLVPKLFYKCDQGOADPVAVARQSSV 959
Db 880 LKQERVLNNVYIHLGDEDPVRHVAASLRLVPKLFYKCDQGOADPVAVARQSSV 939
QY 960 YLKLMLHETQPPSHFSVSTITRIYRGYNLLPSITDVTMNNLSRVIAAVSHELITSTRA 1019
Db 940 YLKLMLHETQPPSHFSVSTITRIYRGYNLLPSITDVTMNNLSRVIAAVSHELITSTRA 999
QY 1020 LTFGCEALCLLSTAFFVCINSLGHGCVPPPLSASDESRSKCTVGMATMILTSSAWFP 1079
Db 1000 LTFGCEALCLLSTAFFVCINSLGHGCVPPPLSASDESRSKCTVGMATMILTSSAWFP 1059
QY 1080 LDLSAHQDALILAGNLLAASAPKLSRSSWASSEANPAATKOEVEWPAALGDRALVPMVEQ 1139
Db 1060 LDLSAHQDALILAGNLLAASAPKLSRSSWASSEANPAATKOEVEWPAALGDRALVPMVEQ 1119
QY 1140 LFSHLLKVINICAHVLDVAPGPAIKAAALPSLTNPSPISPIRRKGKEKEGEGEQASVPLSP 1199
Db 1120 LFSHLLKVINICAHVLDVAPGPAIKAAALPSLTNPSPISPIRRKGKEKEGEGEQASVPLSP 1179
QY 1200 KKGSEASASROSDTSGPVTTSSKSSLSGFSYHPLSYLKLHDVYLKATHANYKVTLDLQNST 1259
Db 1180 KKGSEASASROSDTSGPVTTSSKSSLSGFSYHPLSYLKLHDVYLKATHANYKVTLDLQNST 1239
QY 1260 EKFGFGLRSALDVLSQLLELATLQIDIGKVEEILGLYKSCFSREPMMATVCVQQLLTKLF 1319
Db 1240 EKFGFGLRSALDVLSQLLELATLQIDIGKVEEILGLYKSCFSREPMMATVCVQQLLTKLF 1299
QY 1320 GTNLASQFDGLSSNPKSQGRAQLRGSSSVRPLGYHYCFMAYPTHFTQALADASLRNMVQ 1379
Db 1300 GTNLASQFDGLSSNPKSQGRAQLRGSSSVRPLGYHYCFMAYPTHFTQALADASLRNMVQ 1359
QY 1380 AEQENDTSQFVDVLOKVSQTLKNTLTSVTKNRADKNAIHNIRLPEPLVIKALKQYTTTT 1439
Db 1360 AEQENDTSQFVDVLOKVSQTLKNTLTSVTKNRADKNAIHNIRLPEPLVIKALKQYTTTT 1419
QY 1440 CVQLQKQVLDLALQVLQVRYNCLDSQVDFIGFVLKQFIEYVQFRESAIPNIPFF 1499
Db 1420 SVQLQKQVLDLALQVLQVRYNCLDSQVDFIGFVLKQFIEYVQFRESAIPNIPFF 1479
QY 1500 LVLLSYRHSKQIIGIPKIIQLCDGIMASGRKA 1533
Db 1480 LVLLSYRHSKQIIGIPKIIQLCDGIMASGRKA 1513
```

RESULT 15

PCT-US99-05250-31

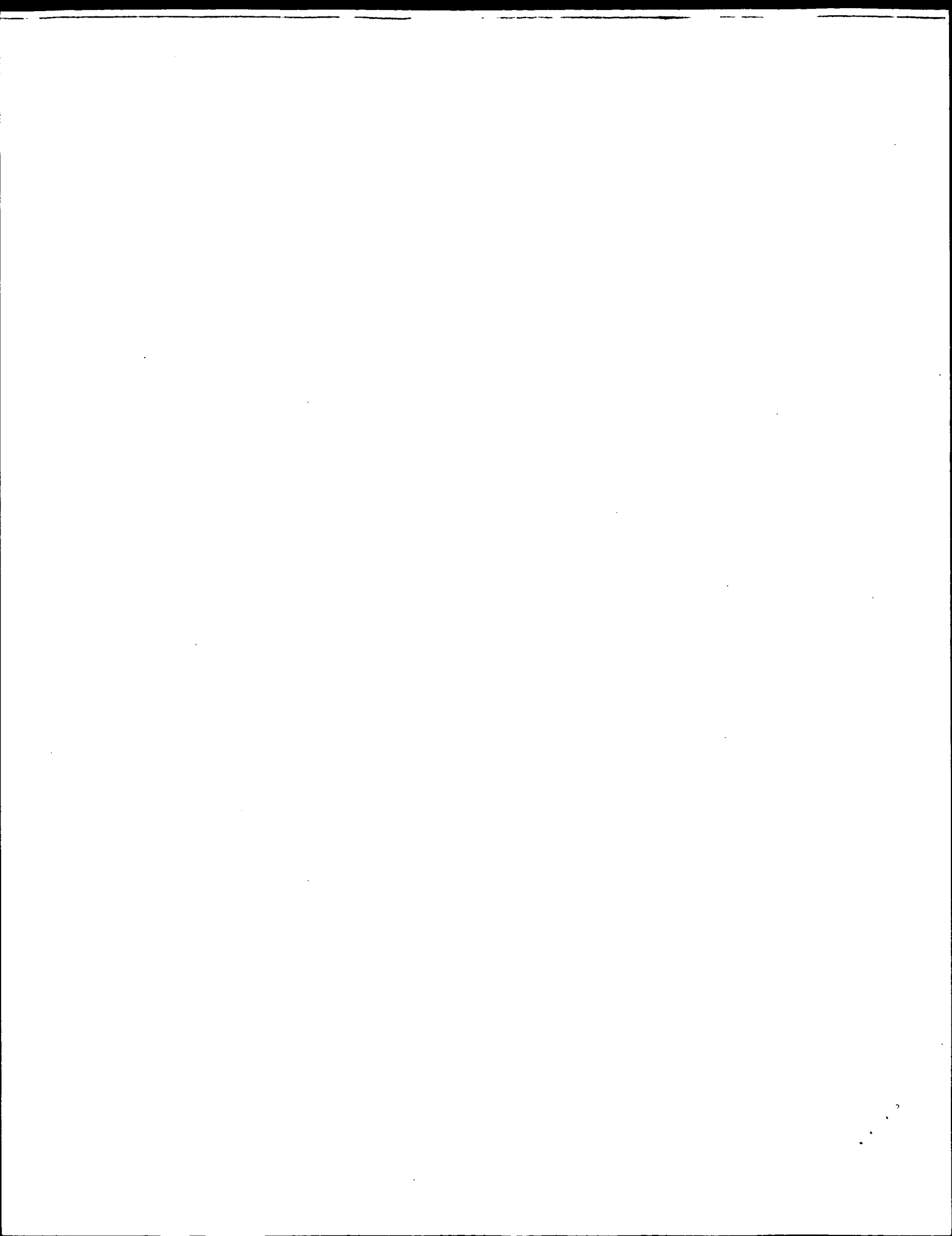
```
; Sequence 31, Application PC/TUS9905250
; GENERAL INFORMATION:
; APPLICANT: The Burnham Institute
; TITLE OF INVENTION: Proapoptotic Peptides, Dependence
; TITLE OF INVENTION: Polypeptides and Methods of Use
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
```

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US99/05250
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/041,886
; FILING DATE: 12-MAR-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: FP-LJ 3484
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 589 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PCT-US99-05250-31
```

```
Query Match 37.6%; Score 2967.5; DB 1; Length 589;
Best Local Similarity 99.5%; Pred. No. 7e-210;
Matches 586; Conservative 0; Mismatches 0; Indels 3; Gaps 1;
```

```
QY 1 MATLEKLMKAFESLKSP---QQQQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPQA 57
Db 1 MATLEKLMKAFESLKSPQQQQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPQA 60
QY 58 QPLLPPQPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 117
Db 61 QPLLPPQPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 120
QY 118 PEFQKLLGIAMELFLLCDDAEDSVRMVADECLNKVIALKALMDSNLPRLQLEYKEIKNG 177
Db 121 PEFQKLLGIAMELFLLCDDAEDSVRMVADECLNKVIALKALMDSNLPRLQLEYKEIKNG 180
QY 178 APRSLRAALWRFELAHVLRPQKCRPYLVNLLPCLTFTSRKRPEESVQETLAAAVPKTMAS 237
Db 181 APRSLRAALWRFELAHVLRPQKCRPYLVNLLPCLTFTSRKRPEESVQETLAAAVPKTMAS 240
QY 238 FGNFANDNEIKVLLKAFIANLKSSSPTRTAAGSAVSICQHSRRTOYFYSWLLNVLGL 297
Db 241 FGNFANDNEIKVLLKAFIANLKSSSPTRTAAGSAVSICQHSRRTOYFYSWLLNVLGL 300
QY 298 LVPVEDEHSTLLILGVLLTLRLVPLVLLQQVQKDTSLKGSFVTRKEMEVSAPSABLQVY 357
Db 301 LVPVEDEHSTLLILGVLLTLRLVPLVLLQQVQKDTSLKGSFVTRKEMEVSAPSABLQVY 360
QY 358 ELTLHHTQHODHNVTGALQLLQQLFRTPPPELLOTLTAVGGIGQLTAAKEESGGRSRG 417
Db 361 ELTLHHTQHODHNVTGALQLLQQLFRTPPPELLOTLTAVGGIGQLTAAKEESGGRSRG 420
QY 418 SIVELIAGGSSCPVLSRKQKGVLLGEBEALDDSESRSDVSSSSALTASVXDEISGEL 477
Db 421 SIVELIAGGSSCPVLSRKQKGVLLGEBEALDDSESRSDVSSSSALTASVXDEISGEL 480
QY 478 AASGVSTPGSAGHDIITEQPRSQHTLQADSVDLASCDLTSSATDGDDEEDILSHSSQVS 537
Db 481 AASGVSTPGSAGHDIITEQPRSQHTLQADSVDLASCDLTSSATDGDDEEDILSHSSQVS 540
QY 538 AVPSDPAMLNDGTQASSPISDSSQTTTEGPDASVTPSDSEIYLDGTD 586
Db 541 AVPSDPAMLNDGTQASSPISDSSQTTTEGPDASVTPSDSEIYLDGTD 589
```

```
Search completed: January 28, 2003, 14:37:35
Job time : 407.975 secs
```



GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:16:13 ; Search time 45.0972 Seconds
(without alignments)
2701.343 Million cell updates/sec

Title: US-09-904-987-7

Perfect score: 7892

Sequence: 1 MATLEKLMKAFESLKSFFQ.....DGIMASGRKASQPYRLCSP 1543

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 394613 seqs, 78952108 residues

Total number of hits satisfying chosen parameters: 394613

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	7821.5	99.1	2182	5	US-09-724-676-51187
2	7821.5	99.1	2182	5	US-09-724-676-51187
3	7821.5	99.1	2324	5	US-09-724-676-51186
4	7821.5	99.1	2324	5	US-09-724-676-51186
5	7821.5	99.1	3223	6	US-10-276-774-1777
6	7821.5	99.1	3223	6	US-10-276-774-1777
7	7086.5	89.8	1433	5	US-09-724-676-51188
8	7086.5	89.8	1433	5	US-09-724-676-51188
9	6008	76.1	1763	5	US-09-724-676-51192
10	6008	76.1	1763	5	US-09-724-676-51192
11	6008	76.1	1940	5	US-09-724-676-51191
12	6008	76.1	1940	5	US-09-724-676-51191
13	5484	69.5	2690	5	US-09-724-676-51190
14	5484	69.5	2690	5	US-09-724-676-51190
15	5273	66.8	1049	5	US-09-724-676-51181
16	5273	66.8	1049	5	US-09-724-676-51181
17	2164.5	27.4	489	5	US-09-724-676-51189
18	2164.5	27.4	489	5	US-09-724-676-51189
19	488	6.2	87	6	US-10-215-432-27
20	373.5	4.7	68	6	US-10-215-432-28
21	200.5	2.5	780	5	US-09-770-689B-5
22	195.5	2.5	159	5	US-09-724-676-60613
23	195.5	2.5	159	5	US-09-724-676-60613
24	195.5	2.5	2496	5	US-09-724-676-91031
25	195.5	2.5	2496	5	US-09-724-676-91031
26	195.5	2.5	2548	5	US-09-724-676-91030

27 195.5 2.5 2548 5 US-09-724-676A-91030 Sequence 91030, A
28 192 2.4 523 6 US-10-017-161-1982 Sequence 1982, Ap
29 192 2.4 1795 6 US-10-170-385-37 Sequence 37, Appl
30 191.5 2.4 2424 5 US-09-724-676-91033 Sequence 91033, A
31 191.5 2.4 2424 5 US-09-724-676A-91033 Sequence 91033, A
32 191.5 2.4 2424 5 US-09-724-676-91032 Sequence 91032, A
33 191.5 2.4 2476 5 US-09-724-676A-91032 Sequence 91032, A
34 191.5 2.4 2731 5 US-09-724-676-91027 Sequence 91027, A
35 191.5 2.4 2731 5 US-09-724-676A-91027 Sequence 91027, A
36 191.5 2.4 2774 5 US-09-724-676-91029 Sequence 91029, A
37 191.5 2.4 2774 5 US-09-724-676A-91029 Sequence 91029, A
38 191.5 2.4 2826 5 US-09-724-676-91028 Sequence 91028, A
39 191.5 2.4 2826 5 US-09-724-676A-91028 Sequence 91028, A
40 190.5 2.4 1104 1 PCT-US02-39133-1 Sequence 1, Appl1
41 184.5 2.3 2328 5 US-09-724-676A-84280 Sequence 84280, A
42 184.5 2.3 2328 5 US-09-724-676A-84280 Sequence 84280, A
43 184.5 2.3 2781 5 US-09-724-676-84279 Sequence 84279, A
44 184.5 2.3 2781 5 US-09-724-676A-84279 Sequence 84279, A
45 180.5 2.3 2346 5 US-09-724-676-84277 Sequence 84277, A

ALIGNMENTS

RESULT 1
US-09-724-676-51187
; Sequence 51187, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724, 676
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 51187
; LENGTH: 2182
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-51187

Query Match 99.1%; Score 7821.5; DB 5; Length 2182;
Best Local Similarity 99.8%; Pred No. 0;
Matches 1533; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 1 MATLEKLMKAFESLKSFF---QQQQQQQQQQQQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPPPPPPPQA 57
DB 43 MATLEKLMKAFESLKSFFQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPPPPPPPQA 102
QY 58 QPLLPOQPP 117
DB 103 QPLLPOQPP 162
QY 118 PFQKLLGIMELFLCSDDAESDVRMVADECLNKVVKALMDSNLPRLQLELYKETKNG 177
DB 163 PFQKLLGIMELFLCSDDAESDVRMVADECLNKVVKALMDSNLPRLQLELYKETKNG 222
QY 178 APRSLRAALWRAELAHVLRPKCRPYLVNLLPCLTRTSKRPEESVQETLAAAVPKIMAS 237
DB 223 APRSLRAALWRAELAHVLRPKCRPYLVNLLPCLTRTSKRPEESVQETLAAAVPKIMAS 282
QY 238 FGNFANDNEIKVLLKAFIANLKSSPTIRTAAGSAVNSICQHSRRTQYFYSWLLNVLLGL 297
DB 283 FGNFANDNEIKVLLKAFIANLKSSPTIRTAAGSAVNSICQHSRRTQYFYSWLLNVLLGL 342
QY 298 LVPVEDEHSTLLILGVLTLRYLVPLLLQQQVQKDTSLKSGFVTRKMEVSPSAEQLVQVY 357
DB 343 LVPVEDEHSTLLILGVLTLRYLVPLLLQQQVQKDTSLKSGFVTRKMEVSPSAEQLVQVY 402
QY 358 ELTLHHTQHODHNVVTGAELELQQLFRTPPPPELLOTLTAVGIGIGOLTAAKESGGRSRG 417
DB 403 ELTLHHTQHODHNVVTGAELELQQLFRTPPPPELLOTLTAVGIGIGOLTAAKESGGRSRG 462

```
QY 418 SIVELTAGGSSCPVLSRKQKGVLLGEEALEDDSESDVSSALTASVKDEISGEL 477
Db 463 SIVELTAGGSSCPVLSRKQKGVLLGEEALEDDSESDVSSALTASVKDEISGEL 522
QY 478 AASSGVSTPGSAGHDIITEQPSQHTLQADSVDLASCDLTSSATDGBEDILSHSSQVS 537
Db 523 AASSGVSTPGSAGHDIITEQPSQHTLQADSVDLASCDLTSSATDGBEDILSHSSQVS 582
QY 538 AVPSDPAMDNDGTQASSPIDSSQTTTEGPDPSAVTPSDSSEIVLDGTDNOYLGIGIQP 597
Db 583 AVPSDPAMDNDGTQASSPIDSSQTTTEGPDPSAVTPSDSSEIVLDGTDNOYLGIGIQP 642
QY 598 QDEDEATGILPDEASEAFRNSSMALQAHLLKNMHCROPSDSSVDKFLVRDEATEPGD 657
Db 643 QDEDEATGILPDEASEAFRNSSMALQAHLLKNMHCROPSDSSVDKFLVRDEATEPGD 702
QY 658 QENKPCRIGDQIGQSTDDSDAPLVHCVRLLSASFLLTGKKNVLPDRDVRVSVKALALSC 717
Db 703 QENKPCRIGDQIGQSTDDSDAPLVHCVRLLSASFLLTGKKNVLPDRDVRVSVKALALSC 762
QY 718 VGAVALHPSEFSEKLYKVPLOTTETPEBOYYSDILNIYIDHGDPOVRGATATLTCGLICS 777
Db 763 VGAVALHPSEFSEKLYKVPLOTTETPEBOYYSDILNIYIDHGDPOVRGATATLTCGLICS 822
QY 778 ILRSRHFVGDWNGTIRTLTGNTFSLADCIPLRLKTLKDESSVTCKLACTAVRNCVMSLC 837
Db 823 ILRSRHFVGDWNGTIRTLTGNTFSLADCIPLRLKTLKDESSVTCKLACTAVRNCVMSLC 882
QY 838 SSSYSELGQLIIDVLTNRSSYWLVRTELLETLAEIDFRLVSFLKAEAKENLHRAHHT 897
Db 883 SSSYSELGQLIIDVLTNRSSYWLVRTELLETLAEIDFRLVSFLKAEAKENLHRAHHT 942
QY 898 GLKQLQERVLNNVYIHLGDEDPVRHVAASLIIRLPKLFYKCDQGOADPVVARQOS 957
Db 943 GLKQLQERVLNNVYIHLGDEDPVRHVAASLIIRLPKLFYKCDQGOADPVVARQOS 1002
QY 958 SVYKLMLHETOPPSHFSYSTIRYRGNLLPSITDVTMNNLSRVIAAASHELITSTT 1017
Db 1003 SVYKLMLHETOPPSHFSYSTIRYRGNLLPSITDVTMNNLSRVIAAASHELITSTT 1062
QY 1018 RALTFCCEALCULLSTAFVPCVLSLWGHGCVPLPSASDESRSKCTVGMATMTLTLSSAW 1077
Db 1063 RALTFCCEALCULLSTAFVPCVLSLWGHGCVPLPSASDESRSKCTVGMATMTLTLSSAW 1122
QY 1078 FPLDLAHDADALLTACNLAAAPKSLRSWASEEENPAATKQEVWPAALGDRLVPVW 1137
Db 1123 FPLDLAHDADALLTACNLAAAPKSLRSWASEEENPAATKQEVWPAALGDRLVPVW 1182
QY 1138 EQLFSLHLKVINTCAHVLDVAPGPAIKAAPSLTNPPSLPIRRKGKEKEPEQASVPL 1197
Db 1183 EQLFSLHLKVINTCAHVLDVAPGPAIKAAPSLTNPPSLPIRRKGKEKEPEQASVPL 1242
QY 1198 SPKKGSEASASQSDTSGPVTTSSKSSLSGSPHYLPSYLLKLDHVLKATHANYKVTLDLQN 1257
Db 1243 SPKKGSEASASQSDTSGPVTTSSKSSLSGSPHYLPSYLLKLDHVLKATHANYKVTLDLQN 1302
QY 1258 STEKFGFLRSALDVLSQLLELATLQDICKVBEILGYLKSCFSREPMMATVCVQQLKT 1317
Db 1303 STEKFGFLRSALDVLSQLLELATLQDICKVBEILGYLKSCFSREPMMATVCVQQLKT 1362
QY 1318 LFGTNLASQPDGLSSNPKSQGRLGSSSVRPGLYHYCFMAYPTHFTQALADASLRNM 1377
Db 1363 LFGTNLASQPDGLSSNPKSQGRLGSSSVRPGLYHYCFMAYPTHFTQALADASLRNM 1422
QY 1378 VQAEQENDTSGWFDVLQKYSTQLKNTLTSVTKNRADKNAIHNHILFELPLVIALKQYTT 1437
Db 1423 VQAEQENDTSGWFDVLQKYSTQLKNTLTSVTKNRADKNAIHNHILFELPLVIALKQYTT 1482
QY 1438 TTCVQLQKQVLDLLAQLVQLRVNYCCLDSQVFIGFVLKQFIEYVQGFRESEAIIPNIF 1497
Db 1483 TTCVQLQKQVLDLLAQLVQLRVNYCCLDSQVFIGFVLKQFIEYVQGFRESEAIIPNIF 1542
QY 1498 FFLVLSYERHSHKQIIGIPKIIQLCDGIMASGRKA 1533
```

```
Db 1543 FFLVLSYERHSHKQIIGIPKIIQLCDGIMASGRKA 1578

RESULT 2
US-09-724-676A-51187
; Sequence 51187, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724, 676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 51187
; LENGTH: 2182
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-51187

Query Match          99.1%; Score 7821.5; DB 5; Length 2182;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1533; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 1 MATLEKLMKAFESLKSF---QQQQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPQA 57
Db 43 MATLEKLMKAFESLKSFQQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPQA 102
QY 58 QPLLPQPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 117
Db 103 QPLLPQPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 162
QY 118 PEFOKLLGIAELFLCDDAESDVRVADENKVIKALMDSNLPRLQLELYKEIKNG 177
Db 163 PEFOKLLGIAELFLCDDAESDVRVADENKVIKALMDSNLPRLQLELYKEIKNG 222
QY 178 APRSLRAALWRFABLAHLVRPQKCRPYLVNLLPCLTRTSKRPEESVOETLAAAVPKIMAS 237
Db 223 APRSLRAALWRFABLAHLVRPQKCRPYLVNLLPCLTRTSKRPEESVOETLAAAVPKIMAS 282
QY 238 FGNFANDNEIKVLLKAFIANLKSSPTIRRTAAGSAVSICOHSRRTQYFYSWLLNVLGL 297
Db 283 FGNFANDNEIKVLLKAFIANLKSSPTIRRTAAGSAVSICOHSRRTQYFYSWLLNVLGL 342
QY 298 LVPVEDEHSTLLILGVLLTLRLYVPLLQQQVKDTSLKGSFGVTRKEMEVSFSAQLVQVY 357
Db 343 LVPVEDEHSTLLILGVLLTLRLYVPLLQQQVKDTSLKGSFGVTRKEMEVSFSAQLVQVY 402
QY 358 ELTLHHTQHDHNVVYTGALQLLQQLFRTPPPPELLQTLTAVGGIGQLTAKEESGGRSRG 417
Db 403 ELTLHHTQHDHNVVYTGALQLLQQLFRTPPPPELLQTLTAVGGIGQLTAKEESGGRSRG 462
QY 418 SIVELIAGGGSSCPVLSRKQKGVLLGEEALEDDSESDVSSALTASVKDEISGEL 477
Db 463 SIVELIAGGGSSCPVLSRKQKGVLLGEEALEDDSESDVSSALTASVKDEISGEL 522
QY 478 AASSGVSTPGSAGHDIITEQPSQHTLQADSVDLASCDLTSSATDGBEDILSHSSQVS 537
Db 523 AASSGVSTPGSAGHDIITEQPSQHTLQADSVDLASCDLTSSATDGBEDILSHSSQVS 582
QY 538 AVPSDPAMDNDGTQASSPIDSSQTTTEGPDPSAVTPSDSSEIVLDGTDNOYLGIGIQP 597
Db 583 AVPSDPAMDNDGTQASSPIDSSQTTTEGPDPSAVTPSDSSEIVLDGTDNOYLGIGIQP 642
QY 598 QDEDEATGILPDEASEAFRNSSMALQAHLLKNMHCROPSDSSVDKFLVRDEATEPGD 657
Db 643 QDEDEATGILPDEASEAFRNSSMALQAHLLKNMHCROPSDSSVDKFLVRDEATEPGD 702
QY 658 QENKPCRIGDQIGQSTDDSDAPLVHCVRLLSASFLLTGKKNVLPDRDVRVSVKALALSC 717
Db 703 QENKPCRIGDQIGQSTDDSDAPLVHCVRLLSASFLLTGKKNVLPDRDVRVSVKALALSC 762
```

QY 718 VGAVALHPESFESKLYKVPDTEPEQYVSDILNIDHGDPOVRGATAILCGTLICS 777
Db 763 VGAVALHPESFESKLYKVPDTEPEQYVSDILNIDHGDPOVRGATAILCGTLICS 822
QY 778 ILSRFRHVGDWMTTIRTLTGNTFSLADCIPLLRKLTLDSESVTKCLACTAVRNCVMSJC 837
Db 823 ILSRFRHVGDWMTTIRTLTGNTFSLADCIPLLRKLTLDSESVTKCLACTAVRNCVMSJC 882
QY 838 SSSYSELGLQIIDLVTLRNSSVWLVRTLETLAEIDFRLVSLFLEAKAENLRGAHHYT 897
Db 883 SSSYSELGLQIIDLVTLRNSSVWLVRTLETLAEIDFRLVSLFLEAKAENLRGAHHYT 942
QY 898 GLKLOERVNNVNIHLLGDEDPVRHVAASLIRLVKPKFYKCDQOGADPVAVARDO 957
Db 943 GLKLOERVNNVNIHLLGDEDPVRHVAASLIRLVKPKFYKCDQOGADPVAVARDO 1002
QY 958 SVYLKLLMHETQPPSHFSVSTIRIYRGYNLLPSITDVTMNNLSRVIAAASHELITST 1017
Db 1003 SVYLKLLMHETQPPSHFSVSTIRIYRGYNLLPSITDVTMNNLSRVIAAASHELITST 1062
QY 1018 RALTFCCEALCLLSTAFPVCINSLGWHGCVPLSLASDESRSKCTVGMATMILTLSSAW 1077
Db 1063 RALTFCCEALCLLSTAFPVCINSLGWHGCVPLSLASDESRSKCTVGMATMILTLSSAW 1122
QY 1078 FPLDLSAHQDALILAGNLLAASAPKSLRSWSASEEANPAATKQEBWVPALGDRLVPMV 1137
Db 1123 FPLDLSAHQDALILAGNLLAASAPKSLRSWSASEEANPAATKQEBWVPALGDRLVPMV 1182
QY 1138 EQLFSLHLKVINICAHVLDVAPGPAIKAALPSLTNPPSLPIRRKGKEPEGEQASVPL 1197
Db 1183 EQLFSLHLKVINICAHVLDVAPGPAIKAALPSLTNPPSLPIRRKGKEPEGEQASVPL 1242
QY 1198 SPKKGSEASASQSDTSGPVTTSKSSLSGSFYHLPYSLKLDHVLKATHANYKVTLDLQ 1257
Db 1243 SPKKGSEASASQSDTSGPVTTSKSSLSGSFYHLPYSLKLDHVLKATHANYKVTLDLQ 1302
QY 1258 STEKFGFLRSALDVLSQLLELATQDQKVEEILGYLKSQFSRPPMMATVCVQQLLKT 1317
Db 1303 STEKFGFLRSALDVLSQLLELATQDQKVEEILGYLKSQFSRPPMMATVCVQQLLKT 1362
QY 1318 LFGTNLASQPDGLSSNPSKQRAQRLGSSSVRPGLYHYCFMAYTHFTQALADASLRNM 1377
Db 1363 LFGTNLASQPDGLSSNPSKQRAQRLGSSSVRPGLYHYCFMAYTHFTQALADASLRNM 1422
QY 1378 VQAEQENDTSGWFDVLQKYSTQLKTNLTSVTKNRADKNAIHNHIRLFEPLVITKALKQYTT 1437
Db 1423 VQAEQENDTSGWFDVLQKYSTQLKTNLTSVTKNRADKNAIHNHIRLFEPLVITKALKQYTT 1482
QY 1438 TTCVQLQKQVLDLLAQLVQLRVNVCLLDSQVFIQFVLFVLFVLFVLFVLFVLFVLFVLFV 1497
Db 1483 TTCVQLQKQVLDLLAQLVQLRVNVCLLDSQVFIQFVLFVLFVLFVLFVLFVLFVLFVLFV 1542
QY 1498 FFLVLLSYERYHSKQIGIPKIIQLCDGIMASGRKA 1533
Db 1543 FFLVLLSYERYHSKQIGIPKIIQLCDGIMASGRKA 1578

RESULT 3

US-09-724-676-51186
; Sequence 51186, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Comugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Comugen
; CURRENT APPLICATION NUMBER: US/09/724, 676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 51186
; LENGTH: 2324
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-51186

Query Match 99.1%; Score 7821.5; DB 5; Length 2324;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1533; Conservative 0; Mismatches 0; Indels 3; Gaps 1;
QY 1 MATLEKMKAFESLKSF---QQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPPPPPQA 57
Db 43 MATLEKMKAFESLKSFQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPPPPPPPQA 102
QY 58 QPLLPPQPP 117
Db 103 QPLLPPQPP 162
QY 118 PEFQKLLGAMFELFLCSDDAESDVRMVADECLNKVIRKALMDSNLPRLQLELYKEIKNG 177
Db 163 PEFQKLLGAMFELFLCSDDAESDVRMVADECLNKVIRKALMDSNLPRLQLELYKEIKNG 222
QY 178 APRSLRAALWRFALHVRPQKCPYLVNLLPCLTRTSKRPEESVQETLAAAVPKIMAS 237
Db 223 APRSLRAALWRFALHVRPQKCPYLVNLLPCLTRTSKRPEESVQETLAAAVPKIMAS 282
QY 238 FGNFANDNEIKVLLKAFIANLKSSSPTIRRTAAGSAVSAICQHSRRTOFYFWLLNVLGL 297
Db 283 FGNFANDNEIKVLLKAFIANLKSSSPTIRRTAAGSAVSAICQHSRRTOFYFWLLNVLGL 342
QY 298 LVPVEDEHSTLLILGVLTLRLVPLLOQVKDTSLKSGFVTRKEMEVSABOLVOVY 357
Db 343 LVPVEDEHSTLLILGVLTLRLVPLLOQVKDTSLKSGFVTRKEMEVSABOLVOVY 402
QY 358 ELTLHHTQHQNHYVTGALELQQLFRTPPPPELLQTLTAVGGIGQLTAAKESGGRSRSG 417
Db 403 ELTLHHTQHQNHYVTGALELQQLFRTPPPPELLQTLTAVGGIGQLTAAKESGGRSRSG 462
QY 418 SIVELIAGGGSCSPVLRRKOKGVLLGEEALEDDSESDVSSSALTASVYKDBISEL 477
Db 463 SIVELIAGGGSCSPVLRRKOKGVLLGEEALEDDSESDVSSSALTASVYKDBISEL 522
QY 478 AASSGVTPGSAGHDITTEPQSOHTLOADSVDLASCDLTSSATDGEDILSHSSQVS 537
Db 523 AASSGVTPGSAGHDITTEPQSOHTLOADSVDLASCDLTSSATDGEDILSHSSQVS 582
QY 538 AVPSDAMDNDGTQASSPISDSQTTTEGPD SATVTPSDSSEIVLDGTDNQTGVLQIQP 597
Db 583 AVPSDAMDNDGTQASSPISDSQTTTEGPD SATVTPSDSSEIVLDGTDNQTGVLQIQP 642
QY 598 QDEDEATGIIPDEASEAFRNSMALQQAHLKNNHSCRPQSDSSVDKFEVLDEATEPGD 657
Db 643 QDEDEATGIIPDEASEAFRNSMALQQAHLKNNHSCRPQSDSSVDKFEVLDEATEPGD 702
QY 658 QENKPCRIGKDIGQSTDDSDAPLVHCVRLLSASFLLITGGKNVLPDRVRVSKALALSC 717
Db 703 QENKPCRIGKDIGQSTDDSDAPLVHCVRLLSASFLLITGGKNVLPDRVRVSKALALSC 762
QY 718 VGAVALHPESFESKLYKVPDTEPEQYVSDILNIDHGDPOVRGATAILCGTLICS 777
Db 763 VGAVALHPESFESKLYKVPDTEPEQYVSDILNIDHGDPOVRGATAILCGTLICS 822
QY 778 ILSRFRHVGDWMTTIRTLTGNTFSLADCIPLLRKLTLDSESVTKCLACTAVRNCVMSJC 837
Db 823 ILSRFRHVGDWMTTIRTLTGNTFSLADCIPLLRKLTLDSESVTKCLACTAVRNCVMSJC 882
QY 838 SSSYSELGLQIIDLVTLRNSSVWLVRTLETLAEIDFRLVSLFLEAKAENLRGAHHYT 897
Db 883 SSSYSELGLQIIDLVTLRNSSVWLVRTLETLAEIDFRLVSLFLEAKAENLRGAHHYT 942
QY 898 GLKLOERVNNVNIHLLGDEDPVRHVAASLIRLVKPKFYKCDQOGADPVAVARDO 957
Db 943 GLKLOERVNNVNIHLLGDEDPVRHVAASLIRLVKPKFYKCDQOGADPVAVARDO 1002
QY 958 SVYLKLLMHETQPPSHFSVSTIRIYRGYNLLPSITDVTMNNLSRVIAAASHELITST 1017
Db 1003 SVYLKLLMHETQPPSHFSVSTIRIYRGYNLLPSITDVTMNNLSRVIAAASHELITST 1062

[illegible]

RESULT 4

```

RESULI* 4
US-09-724-676A-51186
; SEQUENCE 51186, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: CompuGen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 CompuGen
; CURRENT APPLICATION NUMBER: US/09/724.676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 51186
; LENGTH: 2324
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-51186

```

Query Match 99.1%; Score 7821.5; DB 5; Length 2324;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1533; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

[illegible]


```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-51188

Query Match      89.8%; Score 7086.5; DB 5; Length 1433;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1388; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 1 MATLEKMKAFESLSKSF---QQQQQQQQQQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPPPPPQA 57
Db 43 MATLEKMKAFESLSKSFQQQQQQQQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPPPPPQA 102
QY 58 QPLLPOQPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 117
Db 103 QPLLPOQPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 162
QY 118 PFQKLLGTAMLFLLCSDDAEDSVRMVADECLNKVIKALMDSNLPRLQLELYKEIKNG 177
Db 163 PFQKLLGTAMLFLLCSDDAEDSVRMVADECLNKVIKALMDSNLPRLQLELYKEIKNG 222
QY 178 APRSLRAALWRAFAELAHVLRPKCRPYLVNLLPCLTRTSKRPEESVQETLAAAVPKIMAS 237
Db 223 APRSLRAALWRAFAELAHVLRPKCRPYLVNLLPCLTRTSKRPEESVQETLAAAVPKIMAS 282
QY 238 FGNFANDNEIKVLLKAFATANKSSSPTIRRTAAGSAVSIQHSRRTQFYFYSWLLNVLLGL 297
Db 283 FGNFANDNEIKVLLKAFATANKSSSPTIRRTAAGSAVSIQHSRRTQFYFYSWLLNVLLGL 342
QY 298 LVPVEDEHSTLLIGVLLTLRLVPLLOQQVKDTSLKSGFVTRKMEVSPSAEQLVQVY 357
Db 343 LVPVEDEHSTLLIGVLLTLRLVPLLOQQVKDTSLKSGFVTRKMEVSPSAEQLVQVY 402
QY 358 ELTLHHTQHONVVTGALELLQQLFRTPPPELLQTLTAVGGIGQLTAAKESGGRSRG 417
Db 403 ELTLHHTQHONVVTGALELLQQLFRTPPPELLQTLTAVGGIGQLTAAKESGGRSRG 462
QY 418 SIVELIAGGSSCPVLSRKQKGVLLGEEEALEDDESRSDDVSSSALTASVKDEISGEL 477
Db 463 SIVELIAGGSSCPVLSRKQKGVLLGEEEALEDDESRSDDVSSSALTASVKDEISGEL 522
QY 478 AASGVSTPGSAGHDITEQPRSQHTLQADSVDLASCDLTSSATDGEDTLSSSSQVS 537
Db 523 AASGVSTPGSAGHDITEQPRSQHTLQADSVDLASCDLTSSATDGEDTLSSSSQVS 582
QY 538 AVSDPAMDNDGTQASPSIDSSQTTTEGPDSAVTPSDSEIIVLDGTDNQYLGQLGQP 597
Db 583 AVSDPAMDNDGTQASPSIDSSQTTTEGPDSAVTPSDSEIIVLDGTDNQYLGQLGQP 642
QY 598 QDEDEEATGILPDEASEAFNRSSMALQAHLLKNKSHCRQPSDSSVDKFLVRDEATEPGD 657
Db 643 QDEDEEATGILPDEASEAFNRSSMALQAHLLKNKSHCRQPSDSSVDKFLVRDEATEPGD 702
QY 658 QENKPCRIGKIGDSTDDSDAPLVHCVRLLSASFLLTGKKNVLPDRDRVSVKALALSC 717
Db 703 QENKPCRIGKIGDSTDDSDAPLVHCVRLLSASFLLTGKKNVLPDRDRVSVKALALSC 762
QY 718 VGAVALHPESFFSKLYKVPDLTTEPEEQVSDILNYIDHGDPPQVRGATAILCGTLICS 777
Db 763 VGAVALHPESFFSKLYKVPDLTTEPEEQVSDILNYIDHGDPPQVRGATAILCGTLICS 822
QY 778 ILSRSRPHVGDWGMGTIRTLTGNTFSLADCIPLLRKTKLKDSSVTCCKLACTAVRNCVMSLC 837
Db 823 ILSRSRPHVGDWGMGTIRTLTGNTFSLADCIPLLRKTKLKDSSVTCCKLACTAVRNCVMSLC 882
QY 838 SSSYSELGLQLIIDVTLFRNSSYWLVRTELLETLAEIDFRLVSLFLEAKAENLRHGAHHT 897
Db 883 SSSYSELGLQLIIDVTLFRNSSYWLVRTELLETLAEIDFRLVSLFLEAKAENLRHGAHHT 942
QY 898 GLLKQSERVLNNVVIHLLGDEDPVRHVAASLIRLVPKLFYKCDQGDQADPPVAVARDQS 957
Db 943 GLLKQSERVLNNVVIHLLGDEDPVRHVAASLIRLVPKLFYKCDQGDQADPPVAVARDQS 1002
QY 958 SVYLLKLMHETQPPSHFSVSTITRIYRGYNLLPSITDVTMNNLSRVIAAVSHELLITSTT 1017
Db 1003 SVYLLKLMHETQPPSHFSVSTITRIYRGYNLLPSITDVTMNNLSRVIAAVSHELLITSTT 1062
QY 1018 RALTFGCEALCLLSTAFFPVCINSLGWHGCVPLPSASDESRSKCTVGMATMILTLLSSAW 1077
Db 1063 RALTFGCEALCLLSTAFFPVCINSLGWHGCVPLPSASDESRSKCTVGMATMILTLLSSAW 1122
QY 1078 FPLDLSAHQDALIIAGNLLAASAPKSLRSSWASSEEAANPAATKQEEYVWPALGDRALVPMV 1137
Db 1123 FPLDLSAHQDALIIAGNLLAASAPKSLRSSWASSEEAANPAATKQEEYVWPALGDRALVPMV 1182
QY 1138 EQLFSLHLLKVINICAHVLDVAVCPAIPAALPSLTNPSPISPIRRKKEKEPEGEQASVPL 1197
Db 1183 EQLFSLHLLKVINICAHVLDVAVCPAIPAALPSLTNPSPISPIRRKKEKEPEGEQASVPL 1242
QY 1198 SPKKGSEASARSQSDTSGPVTTTSKSSLSGFSYHLPSYKLHDVLYKATHANYKVTLDLQN 1257
Db 1243 SPKKGSEASARSQSDTSGPVTTTSKSSLSGFSYHLPSYKLHDVLYKATHANYKVTLDLQN 1302
QY 1258 STEKFGGFLRSALDVLQIILELATLQDIGKCVBEILGYLYKSCFSREPMMATVCVQQLLKT 1317
Db 1303 STEKFGGFLRSALDVLQIILELATLQDIGKCVBEILGYLYKSCFSREPMMATVCVQQLLKT 1362
QY 1318 LFGTNLASQFDGLSSNPSKSGRAQRUGSSSVRPGLYHYCFMAYPTHFTQALADASLRNM 1377
Db 1363 LFGTNLASQFDGLSSNPSKSGRAQRUGSSSVRPGLYHYCFMAYPTHFTQALADASLRNM 1422
QY 1378 VQAEQENDTSG 1388
Db 1423 VQAEQENDTSG 1433

RESULT 8
US-09-724-676A-51188
; Sequence 51188; Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; NUMBER OF SEQ ID NOS: 2000-11-28
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 51188
; LENGTH: 1433
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-51188

Query Match      89.8%; Score 7086.5; DB 5; Length 1433;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1388; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 1 MATLEKMKAFESLSKSF---QQQQQQQQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPPPPPQA 57
Db 43 MATLEKMKAFESLSKSFQQQQQQQQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPPPPPQA 102
QY 58 QPLLPOQPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 117
Db 103 QPLLPOQPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 162
QY 118 PFQKLLGTAMLFLLCSDDAEDSVRMVADECLNKVIKALMDSNLPRLQLELYKEIKNG 177
Db 163 PFQKLLGTAMLFLLCSDDAEDSVRMVADECLNKVIKALMDSNLPRLQLELYKEIKNG 222
QY 178 APRSLRAALWRAFAELAHVLRPKCRPYLVNLLPCLTRTSKRPEESVQETLAAAVPKIMAS 237
Db 223 APRSLRAALWRAFAELAHVLRPKCRPYLVNLLPCLTRTSKRPEESVQETLAAAVPKIMAS 282
QY 238 FGNFANDNEIKVLLKAFATANKSSSPTIRRTAAGSAVSIQHSRRTQFYFYSWLLNVLLGL 297
Db 283 FGNFANDNEIKVLLKAFATANKSSSPTIRRTAAGSAVSIQHSRRTQFYFYSWLLNVLLGL 342
QY 298 LVPVEDEHSTLLIGVLLTLRLVPLLOQQVKDTSLKSGFVTRKMEVSPSAEQLVQVY 357
Db 343 LVPVEDEHSTLLIGVLLTLRLVPLLOQQVKDTSLKSGFVTRKMEVSPSAEQLVQVY 402
QY 358 ELTLHHTQHONVVTGALELLQQLFRTPPPELLQTLTAVGGIGQLTAAKESGGRSRG 417
Db 403 ELTLHHTQHONVVTGALELLQQLFRTPPPELLQTLTAVGGIGQLTAAKESGGRSRG 462
QY 418 SIVELIAGGSSCPVLSRKQKGVLLGEEEALEDDESRSDDVSSSALTASVKDEISGEL 477
Db 463 SIVELIAGGSSCPVLSRKQKGVLLGEEEALEDDESRSDDVSSSALTASVKDEISGEL 522
QY 478 AASGVSTPGSAGHDITEQPRSQHTLQADSVDLASCDLTSSATDGEDTLSSSSQVS 537
Db 523 AASGVSTPGSAGHDITEQPRSQHTLQADSVDLASCDLTSSATDGEDTLSSSSQVS 582
QY 538 AVSDPAMDNDGTQASPSIDSSQTTTEGPDSAVTPSDSEIIVLDGTDNQYLGQLGQP 597
Db 583 AVSDPAMDNDGTQASPSIDSSQTTTEGPDSAVTPSDSEIIVLDGTDNQYLGQLGQP 642
QY 598 QDEDEEATGILPDEASEAFNRSSMALQAHLLKNKSHCRQPSDSSVDKFLVRDEATEPGD 657
Db 643 QDEDEEATGILPDEASEAFNRSSMALQAHLLKNKSHCRQPSDSSVDKFLVRDEATEPGD 702
QY 658 QENKPCRIGKIGDSTDDSDAPLVHCVRLLSASFLLTGKKNVLPDRDRVSVKALALSC 717
Db 703 QENKPCRIGKIGDSTDDSDAPLVHCVRLLSASFLLTGKKNVLPDRDRVSVKALALSC 762
QY 718 VGAVALHPESFFSKLYKVPDLTTEPEEQVSDILNYIDHGDPPQVRGATAILCGTLICS 777
Db 763 VGAVALHPESFFSKLYKVPDLTTEPEEQVSDILNYIDHGDPPQVRGATAILCGTLICS 822
QY 778 ILSRSRPHVGDWGMGTIRTLTGNTFSLADCIPLLRKTKLKDSSVTCCKLACTAVRNCVMSLC 837
Db 823 ILSRSRPHVGDWGMGTIRTLTGNTFSLADCIPLLRKTKLKDSSVTCCKLACTAVRNCVMSLC 882
QY 838 SSSYSELGLQLIIDVTLFRNSSYWLVRTELLETLAEIDFRLVSLFLEAKAENLRHGAHHT 897
Db 883 SSSYSELGLQLIIDVTLFRNSSYWLVRTELLETLAEIDFRLVSLFLEAKAENLRHGAHHT 942
QY 898 GLLKQSERVLNNVVIHLLGDEDPVRHVAASLIRLVPKLFYKCDQGDQADPPVAVARDQS 957
Db 943 GLLKQSERVLNNVVIHLLGDEDPVRHVAASLIRLVPKLFYKCDQGDQADPPVAVARDQS 1002
QY 958 SVYLLKLMHETQPPSHFSVSTITRIYRGYNLLPSITDVTMNNLSRVIAAVSHELLITSTT 1017
Db 1003 SVYLLKLMHETQPPSHFSVSTITRIYRGYNLLPSITDVTMNNLSRVIAAVSHELLITSTT 1062
QY 1018 RALTFGCEALCLLSTAFFPVCINSLGWHGCVPLPSASDESRSKCTVGMATMILTLLSSAW 1077
Db 1063 RALTFGCEALCLLSTAFFPVCINSLGWHGCVPLPSASDESRSKCTVGMATMILTLLSSAW 1122
QY 1078 FPLDLSAHQDALIIAGNLLAASAPKSLRSSWASSEEAANPAATKQEEYVWPALGDRALVPMV 1137
Db 1123 FPLDLSAHQDALIIAGNLLAASAPKSLRSSWASSEEAANPAATKQEEYVWPALGDRALVPMV 1182
QY 1138 EQLFSLHLLKVINICAHVLDVAVCPAIPAALPSLTNPSPISPIRRKKEKEPEGEQASVPL 1197
Db 1183 EQLFSLHLLKVINICAHVLDVAVCPAIPAALPSLTNPSPISPIRRKKEKEPEGEQASVPL 1242
QY 1198 SPKKGSEASARSQSDTSGPVTTTSKSSLSGFSYHLPSYKLHDVLYKATHANYKVTLDLQN 1257
Db 1243 SPKKGSEASARSQSDTSGPVTTTSKSSLSGFSYHLPSYKLHDVLYKATHANYKVTLDLQN 1302
QY 1258 STEKFGGFLRSALDVLQIILELATLQDIGKCVBEILGYLYKSCFSREPMMATVCVQQLLKT 1317
Db 1303 STEKFGGFLRSALDVLQIILELATLQDIGKCVBEILGYLYKSCFSREPMMATVCVQQLLKT 1362
QY 1318 LFGTNLASQFDGLSSNPSKSGRAQRUGSSSVRPGLYHYCFMAYPTHFTQALADASLRNM 1377
Db 1363 LFGTNLASQFDGLSSNPSKSGRAQRUGSSSVRPGLYHYCFMAYPTHFTQALADASLRNM 1422
QY 1378 VQAEQENDTSG 1388
Db 1423 VQAEQENDTSG 1433
```

```
QY 298 LVPVEDEHSTLLTLGLVLLTLRLVLPLOQVQKDTSLKSGFVTRKEMEVPSPAOLVQVY 357
DB 343 LVPVEDEHSTLLTLGLVLLTLRLVLPLOQVQKDTSLKSGFVTRKEMEVPSPAOLVQVY 402
QY 358 ELTLHTHTQHODHNVTGALQLLQQLFRTPPPELLQTLTAVGGIGQLTAAKESSGRSRG 417
DB 403 ELTLHTHTQHODHNVTGALQLLQQLFRTPPPELLQTLTAVGGIGQLTAAKESSGRSRG 462
QY 418 SIVELIAGGSSCSPVLSRKQKGVLLGEEALEDDSESDVSSSALTASVKDEISGEL 477
DB 463 SIVELIAGGSSCSPVLSRKQKGVLLGEEALEDDSESDVSSSALTASVKDEISGEL 522
QY 478 AASSGVSTPGSAGHDITTEQPSOHTLQADSVDLASCDLTSSATDGEDDILSHSSQVVS 537
DB 523 AASSGVSTPGSAGHDITTEQPSOHTLQADSVDLASCDLTSSATDGEDDILSHSSQVVS 582
QY 538 AVPSDPAAMDNDGTQASSPISDSSTTTTEGPDASVTPSDSSEIVLDGTDNQYLGQIGQP 597
DB 583 AVPSDPAAMDNDGTQASSPISDSSTTTTEGPDASVTPSDSSEIVLDGTDNQYLGQIGQP 642
QY 598 QDEDEEATGILPDEASEAFRNSSMALQQAHLKNMSHCRQPSDSSVDKXVLRDEATEPGD 657
DB 643 QDEDEEATGILPDEASEAFRNSSMALQQAHLKNMSHCRQPSDSSVDKXVLRDEATEPGD 702
QY 658 QENKPCRIGDIGOSTDDDSAPLVHCVRLLSASFLLTGKKNVLPDRDVRVSVKALALSC 717
DB 703 QENKPCRIGDIGOSTDDDSAPLVHCVRLLSASFLLTGKKNVLPDRDVRVSVKALALSC 762
QY 718 VGAVALHPESFFSKLYKVPDTPTEPEQYVSDILNYIDHGPQVVRGATAILCGTFLICS 777
DB 763 VGAVALHPESFFSKLYKVPDTPTEPEQYVSDILNYIDHGPQVVRGATAILCGTFLICS 822
QY 778 ILRSRPHVGDWMTIRTLTGNTFSLADCIPLLRKTLKDESSVTCKLACTAVRNCVMSLC 837
DB 823 ILRSRPHVGDWMTIRTLTGNTFSLADCIPLLRKTLKDESSVTCKLACTAVRNCVMSLC 882
QY 838 SSSYSELGLQIIDLVTLRNSSYWLRTLETLAEIDFRLVFLFLEAKAENLRGAHHYT 897
DB 883 SSSYSELGLQIIDLVTLRNSSYWLRTLETLAEIDFRLVFLFLEAKAENLRGAHHYT 942
QY 898 GLLKQERVLNNVYIHLGDEDPVRHVAAASLIRLVPKLFYKCDQOGADPVVAVARDOS 957
DB 943 GLLKQERVLNNVYIHLGDEDPVRHVAAASLIRLVPKLFYKCDQOGADPVVAVARDOS 1002
QY 958 SVYLKLLMHETQPPSHFSVSTITRIYRGYNLLPSITDVTMENNLSRVIAAVSHELITSTT 1017
DB 1003 SVYLKLLMHETQPPSHFSVSTITRIYRGYNLLPSITDVTMENNLSRVIAAVSHELITSTT 1062
QY 1018 RALTFGCCEALCLLSTAFFVCINSLGWHCGVPPPLSASDESERKSCCTVGMATMILTLLSSAW 1077
DB 1063 RALTFGCCEALCLLSTAFFVCINSLGWHCGVPPPLSASDESERKSCCTVGMATMILTLLSSAW 1122
QY 1078 FPLDLSAQHODALILAGNLLAASAPKSLRSSWASEEENPAATKQEEVWPALGDRALVPMV 1137
DB 1123 FPLDLSAQHODALILAGNLLAASAPKSLRSSWASEEENPAATKQEEVWPALGDRALVPMV 1182
QY 1138 EQLFSLHLLKVINICAHVLDVAPGPAIKAAALPSLTNPSPISPIRRKGEKEPEGEQASVPL 1197
DB 1183 EQLFSLHLLKVINICAHVLDVAPGPAIKAAALPSLTNPSPISPIRRKGEKEPEGEQASVPL 1242
QY 1198 SPKKGSEASASROSQDTSFGPVTTSKSSSLGSFYHLPKSYLKLHDVLRKATHANYKVTLDLQ 1257
DB 1243 SPKKGSEASASROSQDTSFGPVTTSKSSSLGSFYHLPKSYLKLHDVLRKATHANYKVTLDLQ 1302
QY 1258 STEKFGFLRSALDVLQSLILELATLQDITGKCEVEILGYLKSCFSREPMMATVCVQQLLKT 1317
DB 1303 STEKFGFLRSALDVLQSLILELATLQDITGKCEVEILGYLKSCFSREPMMATVCVQQLLKT 1362
QY 1318 LFGTNLASQFDGLSSNPNKSGOGRQIRLSSSVRPGLYHYCFMAYPTHFTQALADASLRNM 1377
DB 1363 LFGTNLASQFDGLSSNPNKSGOGRQIRLSSSVRPGLYHYCFMAYPTHFTQALADASLRNM 1422
QY 1378 VQAEQENDTSG 1388
```

```
DB 1423 VQAEQENDTSG 1433
RESULT 9
US-09-724-676-51192
; Sequence 51192, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Comugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Comugen
; CURRENT APPLICATION NUMBER: US/09/724, 676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 51192
; LENGTH: 1763
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-51192
Query Match 76.1%; Score 6008; DB 5; Length 1763;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1180; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 353 LVQVVELTLHHTQHODHNVTGALQLLQQLFRTPPPELLQTLTAVGGIGQLTAAKESSG 412
DB 14 LFQVVELTLHHTQHODHNVTGALQLLQQLFRTPPPELLQTLTAVGGIGQLTAAKESSG 73
QY 413 RSRGSIIVELTAGGSSCSPVLSRKQKGVLLGEEALEDDSESDVSSSALTASVKDE 472
DB 74 RSRGSIIVELTAGGSSCSPVLSRKQKGVLLGEEALEDDSESDVSSSALTASVKDE 133
QY 473 TSGELAAASGVSTPGSAGHDITTEQPSOHTLQADSVDLASCDLTSSATDGEDDILSHS 532
DB 134 TSGELAAASGVSTPGSAGHDITTEQPSOHTLQADSVDLASCDLTSSATDGEDDILSHS 193
QY 533 SSOVSASVSDPAMDNDGTQASSPISDSSTTTTEGPDASVTPSDSSEIVLDGTDNQYLG 592
DB 194 SSOVSASVSDPAMDNDGTQASSPISDSSTTTTEGPDASVTPSDSSEIVLDGTDNQYLG 253
QY 593 QLGQPDQDEEATGILPDEASEAFRNSSMALQQAHLKNMSHCRQPSDSSVDKXVLRDEA 652
DB 254 QLGQPDQDEEATGILPDEASEAFRNSSMALQQAHLKNMSHCRQPSDSSVDKXVLRDEA 313
QY 653 TEPGQENKPCRIGDIGOSTDDDSAPLVHCVRLLSASFLLTGKKNVLPDRDVRVSVKA 712
DB 314 TEPGQENKPCRIGDIGOSTDDDSAPLVHCVRLLSASFLLTGKKNVLPDRDVRVSVKA 373
QY 713 LALSCVGAVALHPESFFSKLYKVPDTPTEPEQYVSDILNYIDHGPQVVRGATAILCG 772
DB 374 LALSCVGAVALHPESFFSKLYKVPDTPTEPEQYVSDILNYIDHGPQVVRGATAILCG 433
QY 773 TLICSTLSRSRPHVGDWMTIRTLTGNTFSLADCIPLLRKTLKDESSVTCKLACTAVRNC 832
DB 434 TLICSTLSRSRPHVGDWMTIRTLTGNTFSLADCIPLLRKTLKDESSVTCKLACTAVRNC 493
QY 833 VMSLCSYSELGLQIIDLVTLRNSSYWLRTLETLAEIDFRLVFLFLEAKAENLRHG 892
DB 494 VMSLCSYSELGLQIIDLVTLRNSSYWLRTLETLAEIDFRLVFLFLEAKAENLRHG 553
QY 893 AHHTYGLLQERVLNNVYIHLGDEDPVRHVAAASLIRLVPKLFYKCDQOGADPVVAV 952
DB 554 AHHTYGLLQERVLNNVYIHLGDEDPVRHVAAASLIRLVPKLFYKCDQOGADPVVAV 613
QY 953 ARDOSSVYLKLLMHETQPPSHFSVSTITRIYRGYNLLPSITDVTMENNLSRVIAAVSH 1012
DB 614 ARDOSSVYLKLLMHETQPPSHFSVSTITRIYRGYNLLPSITDVTMENNLSRVIAAVSH 673
QY 1013 ITSTTRALTFGCCEALCLLSTAFFVCINSLGWHCGVPPPLSASDESERKSCCTVGMATMILT 1072
DB 674 ITSTTRALTFGCCEALCLLSTAFFVCINSLGWHCGVPPPLSASDESERKSCCTVGMATMILT 733
```

```
QY 1073 LSSAWPDLDSAHQDALILAGNLLAASAPKSLRSSWASEEANPAATKOEWWPALGDRA 1132
D 1132 LSSAWPDLDSAHQDALILAGNLLAASAPKSLRSSWASEEANPAATKOEWWPALGDRA 793
QY 1133 LVPWVEQLFSLHLLKYNINCAHVDVAVGPAKALPSLTNPPSLPIRRKGKEKEPGEQ 1192
D 1192 LVPWVEQLFSLHLLKYNINCAHVDVAVGPAKALPSLTNPPSLPIRRKGKEKEPGEQ 853
QY 1193 ASVPLSPKKGSEASASQSDTSQGVTTSSKSSLGSEFVHLPSYKLHDLVLAATHANYKVT 1252
D 1252 ASVPLSPKKGSEASASQSDTSQGVTTSSKSSLGSEFVHLPSYKLHDLVLAATHANYKVT 913
QY 1253 LDQNSTEKGFLRSALDVLISQILELATLQDQKCVBEIILGYLKSCFSREPMATVCVQ 1312
D 1312 LDQNSTEKGFLRSALDVLISQILELATLQDQKCVBEIILGYLKSCFSREPMATVCVQ 973
QY 1313 QLLKTLFSGTNLASQFDGLSSNPSKSGRAQRLGSSVVRPGLYHYCFMAYTHFTQALADA 1372
D 1372 QLLKTLFSGTNLASQFDGLSSNPSKSGRAQRLGSSVVRPGLYHYCFMAYTHFTQALADA 1033
QY 1373 SLRNMVQAEQENDTSGWFDVLQKYSTQLKTNLTSTVKNRADKNAIHNHRLFEPLVICAL 1432
D 1432 SLRNMVQAEQENDTSGWFDVLQKYSTQLKTNLTSTVKNRADKNAIHNHRLFEPLVICAL 1093
QY 1433 KOYTTTCVQKQKQVLDLQALQVLRVNYCLDSDQVFIGFVLKQFEYIEVQFRESEAI 1492
D 1492 KOYTTTCVQKQKQVLDLQALQVLRVNYCLDSDQVFIGFVLKQFEYIEVQFRESEAI 1153
QY 1493 IPNIFFFVLLSYERYHSKQIIGIPKIIQICDGMASGRKA 1533
D 1533 IPNIFFFVLLSYERYHSKQIIGIPKIIQICDGMASGRKA 1194

RESULT 10
US-09-724-676A-51192
; Sequence 51192, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Comugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Comugen
; CURRENT APPLICATION NUMBER: US/09/724, 676A
; NUMBER OF SEQ ID NOS: 2000-11-28
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 51192
; LENGTH: 1763
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-51192

Query Match 76.1%; Score 6008; DB 5; Length 1763;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1180; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 353 LVQVVELTLLHQTQHDHNVVTGALLELQQLFRTPPPELQTLTAVGIGQLTAKEESGG 412
D 412 LVQVVELTLLHQTQHDHNVVTGALLELQQLFRTPPPELQTLTAVGIGQLTAKEESGG 73
QY 413 RSRSGSIVELIAGGSSCSPVLSRKQKGVLLGEEALEDDSESRDVSALTASVKDE 472
D 472 RSRSGSIVELIAGGSSCSPVLSRKQKGVLLGEEALEDDSESRDVSALTASVKDE 133
QY 473 ISGELASSGVSTPGSAGHDIIITEQPSQHTLQADSVDLASCDLTSSATDGEDDILSHS 532
D 532 ISGELASSGVSTPGSAGHDIIITEQPSQHTLQADSVDLASCDLTSSATDGEDDILSHS 193
QY 533 SSQVSAPVSDPMDLNGTQASSPISDSSQTTTEGPDASVTPDSSSEIVLDGTDNQVGL 592
D 592 SSQVSAPVSDPMDLNGTQASSPISDSSQTTTEGPDASVTPDSSSEIVLDGTDNQVGL 253
QY 593 QIGQPODEDEATGILPDEASEAFRNSMALQQAHLKNNSHCRQPSDSSVDKFLRDEA 652
D 652 QIGQPODEDEATGILPDEASEAFRNSMALQQAHLKNNSHCRQPSDSSVDKFLRDEA 313
```

```
QY 653 TEPGDQENKPCRIKGDITGQSTDDDSAPLVHCVRLLSASFLLTGKKNVLPDROVRVSVKA 712
D 712 TEPGDQENKPCRIKGDITGQSTDDDSAPLVHCVRLLSASFLLTGKKNVLPDROVRVSVKA 373
QY 713 LALSCVGAVALHPESFESKLYKVPDLDTTEYPERQYYSIDLINLYDHGDPQVRGATILCG 772
D 772 LALSCVGAVALHPESFESKLYKVPDLDTTEYPERQYYSIDLINLYDHGDPQVRGATILCG 433
QY 773 TLICSIILSRFRHVGDMGTIRTLTGNTFTSLADCIPLLRKTKLDESSVTCKLACTAVRNC 832
D 832 TLICSIILSRFRHVGDMGTIRTLTGNTFTSLADCIPLLRKTKLDESSVTCKLACTAVRNC 493
QY 833 VMSLCSSSYSELGLQIILIDVLTLRNSSYWLVRTELLETLAEIDFRLVSLFLEAKAENLHRG 892
D 892 VMSLCSSSYSELGLQIILIDVLTLRNSSYWLVRTELLETLAEIDFRLVSLFLEAKAENLHRG 553
QY 893 AHHTYGLKQLQERVNNVVIHLLGDEDPVRHVAASLIRLVPKLFYKCDQGOADPVAV 952
D 952 AHHTYGLKQLQERVNNVVIHLLGDEDPVRHVAASLIRLVPKLFYKCDQGOADPVAV 613
QY 953 ARDQSSVYLKLMHETQPPSHFSVSTITRIYRGYNLLPSITDVTMNNLSRVIAAVSHEL 1012
D 1012 ARDQSSVYLKLMHETQPPSHFSVSTITRIYRGYNLLPSITDVTMNNLSRVIAAVSHEL 673
QY 1013 ITSTTRALTFCGCEALCLLSTAPVCIWSLGHGCVPPPLSASDESRSKSTVGMATMILT 1072
D 1072 ITSTTRALTFCGCEALCLLSTAPVCIWSLGHGCVPPPLSASDESRSKSTVGMATMILT 733
QY 1073 LSSAWPDLDSAHQDALILAGNLLAASAPKSLRSSWASEEANPAATKOEWWPALGDRA 1132
D 1132 LSSAWPDLDSAHQDALILAGNLLAASAPKSLRSSWASEEANPAATKOEWWPALGDRA 793
QY 1133 LVPWVEQLFSLHLLKYNINCAHVDVAVGPAKALPSLTNPPSLPIRRKGKEKEPGEQ 1192
D 1192 LVPWVEQLFSLHLLKYNINCAHVDVAVGPAKALPSLTNPPSLPIRRKGKEKEPGEQ 853
QY 1193 ASVPLSPKKGSEASASQSDTSQGVTTSSKSSLGSEFVHLPSYKLHDLVLAATHANYKVT 1252
D 1252 ASVPLSPKKGSEASASQSDTSQGVTTSSKSSLGSEFVHLPSYKLHDLVLAATHANYKVT 913
QY 1253 LDQNSTEKGFLRSALDVLISQILELATLQDQKCVBEIILGYLKSCFSREPMATVCVQ 1312
D 1312 LDQNSTEKGFLRSALDVLISQILELATLQDQKCVBEIILGYLKSCFSREPMATVCVQ 973
QY 1313 QLLKTLFSGTNLASQFDGLSSNPSKSGRAQRLGSSVVRPGLYHYCFMAYTHFTQALADA 1372
D 1372 QLLKTLFSGTNLASQFDGLSSNPSKSGRAQRLGSSVVRPGLYHYCFMAYTHFTQALADA 1033
QY 1373 SLRNMVQAEQENDTSGWFDVLQKYSTQLKTNLTSTVKNRADKNAIHNHRLFEPLVICAL 1432
D 1432 SLRNMVQAEQENDTSGWFDVLQKYSTQLKTNLTSTVKNRADKNAIHNHRLFEPLVICAL 1093
QY 1433 KOYTTTCVQKQKQVLDLQALQVLRVNYCLDSDQVFIGFVLKQFEYIEVQFRESEAI 1492
D 1492 KOYTTTCVQKQKQVLDLQALQVLRVNYCLDSDQVFIGFVLKQFEYIEVQFRESEAI 1153
QY 1493 IPNIFFFVLLSYERYHSKQIIGIPKIIQICDGMASGRKA 1533
D 1533 IPNIFFFVLLSYERYHSKQIIGIPKIIQICDGMASGRKA 1194

RESULT 11
US-09-724-676-51191
; Sequence 51191, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Comugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Comugen
; CURRENT APPLICATION NUMBER: US/09/724, 676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
```

; SEQ ID NO 51191
; LENGTH: 1940
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-51191

Query Match 76.1%; Score 6008; DB 5; Length 1940;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1180; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 353 LVQVYELTTHHTQHODHNVTGALLELQQLFRTPPPELLQTLTAVGGIGQLTAAKEESGG 412
Db 14 LFQVYELTTHHTQHODHNVTGALLELQQLFRTPPPELLQTLTAVGGIGQLTAAKEESGG 73
QY 413 RSRGSGIVELIAGGSSCPVLSRKQKGVLLGEEALEDDESDSSVDFVLRDEA 472
Db 74 RSRGSGIVELIAGGSSCPVLSRKQKGVLLGEEALEDDESDSSVDFVLRDEA 133
QY 473 ISGELAASSGVSTPGSAGHDITTEQPRSOHTLQADSVDLASCDLTSSATDGEEDILSHS 532
Db 134 ISGELAASSGVSTPGSAGHDITTEQPRSOHTLQADSVDLASCDLTSSATDGEEDILSHS 193
QY 533 SSOVSAPSDPAMDNDGTQASSPISDSSTQTTTEGPDSSAVTPSDSSEIVLDGTNOYLGL 592
Db 194 SSOVSAPSDPAMDNDGTQASSPISDSSTQTTTEGPDSSAVTPSDSSEIVLDGTNOYLGL 253
QY 593 QIQOPODEEATGILPDEASEAFRNSMALQOAHLLKNMHCROPSDSSVDKFLRDEA 652
Db 254 QIQOPODEEATGILPDEASEAFRNSMALQOAHLLKNMHCROPSDSSVDKFLRDEA 313
QY 653 TEPGQDNKPCRIGKDGIGOSTDDSDAPLVHCVRLLSASFLLTGKKNVLPDRDVRVSKA 712
Db 314 TEPGQDNKPCRIGKDGIGOSTDDSDAPLVHCVRLLSASFLLTGKKNVLPDRDVRVSKA 373
QY 713 LALSCVGAVALHPESFFSKLVPLDTTEYBEEQVSDILNYIDHGDPOVRGATAILCG 772
Db 374 LALSCVGAVALHPESFFSKLVPLDTTEYBEEQVSDILNYIDHGDPOVRGATAILCG 433
QY 773 TLICSIILSRSRPHVGMGTIRTLTGNTFSLADCIPLLRKTLKDESSVTCCKLACTAVRNC 832
Db 434 TLICSIILSRSRPHVGMGTIRTLTGNTFSLADCIPLLRKTLKDESSVTCCKLACTAVRNC 553
QY 833 VMSLCSSSYSELGQLIIDVLTLRNSSYWLVRTELLETLAEIDFRVLSFEAKAENLHRG 892
Db 494 VMSLCSSSYSELGQLIIDVLTLRNSSYWLVRTELLETLAEIDFRVLSFEAKAENLHRG 553
QY 893 AHHTGLLKQERVLNNVYIHLGDEDPVRVHVAASLRLVPLKFKYCDQGOADPVAV 952
Db 554 AHHTGLLKQERVLNNVYIHLGDEDPVRVHVAASLRLVPLKFKYCDQGOADPVAV 613
QY 953 ARQSSVYLKLLMHETOPPSHFVSITIRYGYNLLPSITDVTMNNLSRVIAAVSHSL 1012
Db 614 ARQSSVYLKLLMHETOPPSHFVSITIRYGYNLLPSITDVTMNNLSRVIAAVSHSL 673
QY 1013 ITSTTRALTFCCEALCLLSTAFPVCIVSLGWHCGVPPPLSASDESRSKCTVGMATMILTL 1072
Db 674 ITSTTRALTFCCEALCLLSTAFPVCIVSLGWHCGVPPPLSASDESRSKCTVGMATMILTL 733
QY 1073 LSSANFPDLISAQDALILAGNLLAASAPKSLRSSWASEEANPAATKQEEVWPALGDRA 1132
Db 734 LSSANFPDLISAQDALILAGNLLAASAPKSLRSSWASEEANPAATKQEEVWPALGDRA 793
QY 1133 LVPVVEQLFSLHLKVINICAHVLDVAPGAIPAALPSLTNNPSPISPIRRKKEKEPGEQ 1192
Db 794 LVPVVEQLFSLHLKVINICAHVLDVAPGAIPAALPSLTNNPSPISPIRRKKEKEPGEQ 853
QY 1193 ASVPLSPKKGSEASASRQSDTSGPVVTTSSKSLGSFYHLPSYKLKLDHVLKATHYKVT 1252
Db 854 ASVPLSPKKGSEASASRQSDTSGPVVTTSSKSLGSFYHLPSYKLKLDHVLKATHYKVT 913
QY 1253 LDLONSTEKGGLRSALDVLSQLILELATLQDYGKCVETLGYLKSCEPREPMATVCVQ 1312
Db 914 LDLONSTEKGGLRSALDVLSQLILELATLQDYGKCVETLGYLKSCEPREPMATVCVQ 973
```

```
QY 1313 QLLKTLFCTNLASQFDGLSSNPSSKSGRAQRLGSSSVRPGLYHYCFMAPYTHFTQALADA 1372
Db 974 QLLKTLFCTNLASQFDGLSSNPSSKSGRAQRLGSSSVRPGLYHYCFMAPYTHFTQALADA 1033
QY 1373 SLRNNVQAEQENDTSGWFDVLQKVSTOLKTNLTSVTKNRADKNAIHNHILRFLPELVKAL 1432
Db 1034 SLRNNVQAEQENDTSGWFDVLQKVSTOLKTNLTSVTKNRADKNAIHNHILRFLPELVKAL 1093
QY 1433 KOYTTTTCVOLQKOVLDLALQVLQVRVNYCULDSQVFIGFVLKQFYEIVGQFRESEAI 1492
Db 1094 KOYTTTTCVOLQKOVLDLALQVLQVRVNYCULDSQVFIGFVLKQFYEIVGQFRESEAI 1153
QY 1493 IPNIFFFVLVLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1533
Db 1154 IPNIFFFVLVLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1194
```

RESULT 12

US-09-724-676A-51191

; Sequence 51191, Application us/09724676A

; GENERAL INFORMATION:

; APPLICANT: Compugen LTD

; TITLE OF INVENTION: Variants of alternative splicing

; FILE REFERENCE: 129181.4 Compugen

; CURRENT APPLICATION NUMBER: US/09/724,676A

; CURRENT FILING DATE: 2000-11-28

; NUMBER OF SEQ ID NOS: 97222

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 51191

; LENGTH: 1940

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-724-676A-51191

Query Match 76.1%; Score 6008; DB 5; Length 1940;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1180; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 353 LVQVYELTTHHTQHODHNVTGALLELQQLFRTPPPELLQTLTAVGGIGQLTAAKEESGG 412
Db 14 LFQVYELTTHHTQHODHNVTGALLELQQLFRTPPPELLQTLTAVGGIGQLTAAKEESGG 73
QY 413 RSRGSGIVELIAGGSSCPVLSRKQKGVLLGEEALEDDESDSSVDFVLRDEA 472
Db 74 RSRGSGIVELIAGGSSCPVLSRKQKGVLLGEEALEDDESDSSVDFVLRDEA 133
QY 473 ISGELAASSGVSTPGSAGHDITTEQPRSOHTLQADSVDLASCDLTSSATDGEEDILSHS 532
Db 134 ISGELAASSGVSTPGSAGHDITTEQPRSOHTLQADSVDLASCDLTSSATDGEEDILSHS 193
QY 533 SSOVSAPSDPAMDNDGTQASSPISDSSTQTTTEGPDSSAVTPSDSSEIVLDGTNOYLGL 592
Db 194 SSOVSAPSDPAMDNDGTQASSPISDSSTQTTTEGPDSSAVTPSDSSEIVLDGTNOYLGL 253
QY 593 QIQOPODEEATGILPDEASEAFRNSMALQOAHLLKNMHCROPSDSSVDKFLRDEA 652
Db 254 QIQOPODEEATGILPDEASEAFRNSMALQOAHLLKNMHCROPSDSSVDKFLRDEA 313
QY 653 TEPGQDNKPCRIGKDGIGOSTDDSDAPLVHCVRLLSASFLLTGKKNVLPDRDVRVSKA 712
Db 314 TEPGQDNKPCRIGKDGIGOSTDDSDAPLVHCVRLLSASFLLTGKKNVLPDRDVRVSKA 373
QY 713 LALSCVGAVALHPESFFSKLVPLDTTEYBEEQVSDILNYIDHGDPOVRGATAILCG 772
Db 374 LALSCVGAVALHPESFFSKLVPLDTTEYBEEQVSDILNYIDHGDPOVRGATAILCG 433
QY 773 TLICSIILSRSRPHVGMGTIRTLTGNTFSLADCIPLLRKTLKDESSVTCCKLACTAVRNC 832
Db 434 TLICSIILSRSRPHVGMGTIRTLTGNTFSLADCIPLLRKTLKDESSVTCCKLACTAVRNC 493
QY 833 VMSLCSSSYSELGQLIIDVLTLRNSSYWLVRTELLETLAEIDFRVLSFEAKAENLHRG 892
```

Db 494 VMSLCSSSYSELGLQLIIDVLTNRSSYWLVRTELETLAEIDPRLVSFLEAKAENLHRG 553
QY 893 AHHTYGLKLOERVLNNVVIHLLGDEDRVRHVAASLIRLPKLFYKCDGQADPVVAV 952
Db 554 AHHTYGLKLOERVLNNVVIHLLGDEDRVRHVAASLIRLPKLFYKCDGQADPVVAV 613
QY 953 ARDQSSVYLKLMHETQPPSHFSVSTIIRYRGYNLLPSITDVTMNNLSRVIAAVSHL 1012
Db 614 ARDQSSVYLKLMHETQPPSHFSVSTIIRYRGYNLLPSITDVTMNNLSRVIAAVSHL 673
QY 1013 ITSTTRALTFCGCEALCLLSTAFPVCIVSLGWHCGVPPLSASDESRSKCTVGMATMILT 1072
Db 674 ITSTTRALTFCGCEALCLLSTAFPVCIVSLGWHCGVPPLSASDESRSKCTVGMATMILT 733
QY 1073 LSSANFPDLSSAHQDALILAGNLLAASAPKSLRSSWASEEENPAATKQEEVWALGDRA 1132
Db 734 LSSANFPDLSSAHQDALILAGNLLAASAPKSLRSSWASEEENPAATKQEEVWALGDRA 793
QY 1133 LVPVMEQLFSLHLLKVINICAHVLDVAVPAPAIKAALPSLTNPPSLSPRRKGKEPEGEQ 1192
Db 794 LVPVMEQLFSLHLLKVINICAHVLDVAVPAPAIKAALPSLTNPPSLSPRRKGKEPEGEQ 853
QY 1193 ASVPLSPKKGSEASASRSQSDTSGPVTTSKSSLSGSFYHLPSYKLHDVVKATHANYKVT 1252
Db 854 ASVPLSPKKGSEASASRSQSDTSGPVTTSKSSLSGSFYHLPSYKLHDVVKATHANYKVT 913
QY 1253 LDQNSTEKGFGFLRSALDVLSQLLELATLQDQKVEEILGYLKSCFSREPMMATVCVQ 1312
Db 914 LDQNSTEKGFGFLRSALDVLSQLLELATLQDQKVEEILGYLKSCFSREPMMATVCVQ 973
QY 1313 QLLKTLFTCNLASQFDGLSSNPSKSGRAQLGSSSVRPGLYHYCFMAYPTHFTQALADA 1372
Db 974 QLLKTLFTCNLASQFDGLSSNPSKSGRAQLGSSSVRPGLYHYCFMAYPTHFTQALADA 1033
QY 1373 SLRNVMQAEQNDTSGWFDVLQKVSQTLKNTLTSVTKNRADKNATHIRLFEPLVIKAL 1432
Db 1034 SLRNVMQAEQNDTSGWFDVLQKVSQTLKNTLTSVTKNRADKNATHIRLFEPLVIKAL 1093
QY 1433 QYTTTTTCVQKQKQVLDLQALVOLRVNVCLLDSQVDFIGVLFKQFEVIEVQFSEAI 1492
Db 1094 QYTTTTTCVQKQKQVLDLQALVOLRVNVCLLDSQVDFIGVLFKQFEVIEVQFSEAI 1153
QY 1493 IPNIFFLVLLSYERYHSHKQIIGIPKIIQLCDGIMASGRKA 1533
Db 1154 IPNIFFLVLLSYERYHSHKQIIGIPKIIQLCDGIMASGRKA 1194

RESULT 13
US-09-724-676-51190
; Sequence 51190, Application us/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724.676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 51190
; LENGTH: 2690
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-51190

Query Match 69.5%; Score 5484; DB 5; Length 2690;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1075; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 459 DVSSSALTASVKDEISGELAASSGVSTPGSAGHDIIITEQPSQHTLOADSVLASCDDLTS 518
Db 8 DVSSSALTASVKDEISGELAASSGVSTPGSAGHDIIITEQPSQHTLOADSVLASCDDLTS 67
QY 519 SATDGEDEEDILSHSSSQVSAVPSDPAMDNDGTQASSPISDSSQTTTEGPDPSAVTPSDSS 578

Db 68 SATDGEDEEDILSHSSSQVSAVPSDPAMDNDGTQASSPISDSSQTTTEGPDPSAVTPSDSS 127
QY 579 EIVLDGTDNOYLGIGIQPODEDEATGILPDEASEAFRNSMALQQAHLKNMSHCRQP 638
Db 128 EIVLDGTDNOYLGIGIQPODEDEATGILPDEASEAFRNSMALQQAHLKNMSHCRQP 187
QY 639 SDSSVDFVLRDEATEPGDQENPCRIKGDIGQSTDDDSAPLVHCVRLLSASFLLTGKGN 698
Db 188 SDSSVDFVLRDEATEPGDQENPCRIKGDIGQSTDDDSAPLVHCVRLLSASFLLTGKGN 247
QY 699 VLPVDRVRYSKALALUSCVGAVALHPESFKLYKVPDQTTTEYPEQVVSIDLNVYIDH 758
Db 248 VLPVDRVRYSKALALUSCVGAVALHPESFKLYKVPDQTTTEYPEQVVSIDLNVYIDH 307
QY 759 GDPQVRGATAILCGTLLCSILSRSHRPHVGMGTIRTLTGNTFSLADICPLLRKTKLDES 818
Db 308 GDPQVRGATAILCGTLLCSILSRSHRPHVGMGTIRTLTGNTFSLADICPLLRKTKLDES 367
QY 819 SVTCKLACTAVRNCVMSLCSSSYSELGLQLIIDVLTNRSSYWLVRTELETLAEIDFRL 878
Db 368 SVTCKLACTAVRNCVMSLCSSSYSELGLQLIIDVLTNRSSYWLVRTELETLAEIDFRL 427
QY 879 VSFEAKAENLHRGAHHYTGLLKLOERVLNNVVIHLLGDEDRVRHVAASLIRLPKLF 938
Db 428 VSFEAKAENLHRGAHHYTGLLKLOERVLNNVVIHLLGDEDRVRHVAASLIRLPKLF 487
QY 939 KYCDQGDQADPVAVARDQSSVYLKLMHETQPPSHFSVSTIIRYRGYNLLPSITDVTME 998
Db 488 KYCDQGDQADPVAVARDQSSVYLKLMHETQPPSHFSVSTIIRYRGYNLLPSITDVTME 547
QY 999 NNLSRVTAAYSHELITSTTRALTFCGCEALCLLSTAPPVCIVSLGWHCGVPPLSASDES 1058
Db 548 NNLSRVTAAYSHELITSTTRALTFCGCEALCLLSTAPPVCIVSLGWHCGVPPLSASDES 607
QY 1059 KSCVTGNATMILTLLSSANFPDLSSAHQDALILAGNLLAASAPKSLRSSWASEEENPA 1118
Db 608 KSCVTGNATMILTLLSSANFPDLSSAHQDALILAGNLLAASAPKSLRSSWASEEENPA 667
QY 1119 TKQEEVWALGDRAVPMVEQLFSLHLLKVINICAHVLDVAVPAPAIKAALPSLTNPPSL 1178
Db 668 TKQEEVWALGDRAVPMVEQLFSLHLLKVINICAHVLDVAVPAPAIKAALPSLTNPPSL 727
QY 1179 PIRKKEKEPEGEQASVPLSPKKGSEASASRSQSDTSGPVTTSKSSLSGSFYHLPSYKL 1238
Db 728 PIRKKEKEPEGEQASVPLSPKKGSEASASRSQSDTSGPVTTSKSSLSGSFYHLPSYKL 787
QY 1239 HDVLKATHANYKVTLDLQNSTEKGFGFLRSALDVLSQLLELATLQDQKVEEILGYLKS 1298
Db 788 HDVLKATHANYKVTLDLQNSTEKGFGFLRSALDVLSQLLELATLQDQKVEEILGYLKS 847
QY 1299 CFSREPMMATVCVQQLKTLFTGNLASQFDGLSSNPSKSGRAQLGSSSVRPGLYHYCF 1358
Db 848 CFSREPMMATVCVQQLKTLFTGNLASQFDGLSSNPSKSGRAQLGSSSVRPGLYHYCF 907
QY 1359 MAPYTHFTQALADASLRNMVQAEQNDTSGWFDVLQKVSQTLKNTLTSVTKNRADKNATH 1418
Db 908 MAPYTHFTQALADASLRNMVQAEQNDTSGWFDVLQKVSQTLKNTLTSVTKNRADKNATH 967
QY 1419 NHIRLFEPLVIKALQYTTTTTCVQKQKQVLDLQALVOLRVNVCLLDSQVDFIGVLFKQF 1478
Db 968 NHIRLFEPLVIKALQYTTTTTCVQKQKQVLDLQALVOLRVNVCLLDSQVDFIGVLFKQF 1027
QY 1479 EYIEVQFSEAIIPNIFFLVLLSYERYHSHKQIIGIPKIIQLCDGIMASGRKA 1533
Db 1028 EYIEVQFSEAIIPNIFFLVLLSYERYHSHKQIIGIPKIIQLCDGIMASGRKA 1082

RESULT 14
US-09-724-676A-51190
; Sequence 51190, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD


```

; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724_676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 51190
; LENGTH: 2690
; TYPE: prt
; ORGANISM: Homo sapiens
; US-09-724-676A-51190

```

```
Query Match          69.5%; Score 5484; DB 5; Length 2690;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1075; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy	459	DVSSALTASVKDEISGEIAAASGVSTPGSAGHDIIITEQPRSQHTLQADSVDLASCDLTS	518
Db	8	DVSSALTASVKDEISGEIAAASGVSTPGSAGHDIIITEQPRSQHTLQADSVDLASCDLTS	67
Qy	519	SATDGDDEEDTLSSHSSQVSAVPSDPAAMDNDGTOASSPISDSSQTTTEGPDSAVTPSDSS	578
Db	68	SATDGDDEEDTLSSHSSQVSAVPSDPAAMDNDGTOASSPISDSSQTTTEGPDSAVTPSDSS	127
Qy	579	EIVLDGTDNQYLGQIQCPQDEDEEATGILPDEASEAFRNSMALQQAHLTKNMHCRQP	638
Db	128	EIVLDGTDNQYLGQIQCPQDEDEEATGILPDEASEAFRNSMALQQAHLTKNMHCRQP	187
Qy	639	SDSSVDKFVLRDSEATPGDQENKPCRIKGDIGQSTDDSDAPLVHCVRLLSASFLLTGCKN	698
Db	188	SDSSVDKFVLRDSEATPGDQENKPCRIKGDIGQSTDDSDAPLVHCVRLLSASFLLTGCKN	247
Qy	699	VLVPDRDVRVSKALASCVGAAVLHPESFESKLYKVPLOTTTEYPBEOQYVSDILNTIDH	758
Db	248	VLVPDRDVRVSKALASCVGAAVLHPESFESKLYKVPLOTTTEYPBEOQYVSDILNTIDH	307
Qy	759	GDPOVRGATAILCGTILCSILSRSRPHVGDMWGTTRTITGNTFSLADCIPLLRKTLKDES	818
Db	308	GDPOVRGATAILCGTILCSILSRSRPHVGDMWGTTRTITGNTFSLADCIPLLRKTLKDES	367
Qy	819	SVTCKLACTAVRCNVMSLCSSSVSELGLQLIIDVLTLRNSSVWLVRTELETTLAEIDFRL	878
Db	368	SVTCKLACTAVRCNVMSLCSSSVSELGLQLIIDVLTLRNSSVWLVRTELETTLAEIDFRL	427
Qy	879	VSFLEAKAENLHRGAHHYTGLLKLQERVUNNNVYIHLGDEDPVRVRHVAASLIRLVPKLF	938
Db	428	VSFLEAKAENLHRGAHHYTGLLKLQERVUNNNVYIHLGDEDPVRVRHVAASLIRLVPKLF	487
Qy	939	YKCDQGAQDPVAVARDQSSVYLKLLMHETQPPSHFSVSTIRIYRGYNLLPSTIDVTME	998
Db	488	YKCDQGAQDPVAVARDQSSVYLKLLMHETQPPSHFSVSTIRIYRGYNLLPSTIDVTME	547
Qy	999	NNLSRVIAAVSHELITSTTRALTFCCEALCLLSTAFVPCIVSLGWCHGCVPLSASDES	1058
Db	548	NNLSRVIAAVSHELITSTTRALTFCCEALCLLSTAFVPCIVSLGWCHGCVPLSASDES	607
Qy	1059	KSCTVGMATMILPLLSSANFPDLSHAHQDALILAGNLLAASAPKSLRSSWSASEEANPAA	1118
Db	608	KSCTVGMATMILPLLSSANFPDLSHAHQDALILAGNLLAASAPKSLRSSWSASEEANPAA	667
Qy	1119	TQBEVWPALCDRALVPWVEOLFSLHLLKVINICAHVLDVDPAGPAIKAALPSLNPSPSL	1178
Db	668	TQBEVWPALCDRALVPWVEOLFSLHLLKVINICAHVLDVDPAGPAIKAALPSLNPSPSL	727
Qy	1179	PIRKKGKEPGEQAQVYPLSPKKGSEASARSQSDTSGPVTITSKSSSLGSGYHLPYSYKL	1238
Db	728	PIRKKGKEPGEQAQVYPLSPKKGSEASARSQSDTSGPVTITSKSSSLGSGYHLPYSYKL	787
Qy	1239	HDVLKATHANYKVTLDLONSTKFGFGLRSADVLQIILELATLQDICKVCVEETILGYLKS	1298
Db	788	HDVLKATHANYKVTLDLONSTKFGFGLRSADVLQIILELATLQDICKVCVEETILGYLKS	847
Qy	1299	CF\$REPMMATVVCVQQLKTLTFLGTNLA\$OPQDGLSSNPNK\$SQG\$RAQL\$GLSSVRPGLYHCF	1358

```

RESULT 15
US-09-724-676-51181
; Sequence 51181, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724.676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 51181
; LENGTH: 1049
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-51181

```

Query Match	66.8%	Score 5273;	DB 5;	Length 1049;
Best Local Similarity	99.9%;	Pred. No. 0;		
Matches 1035: Conservative	0;	Mismatches	1;	Indels 0; Gaps 0;

Qy	353	LVQYELTLHRTQHQDHNVVTG	GALELLQQFLRTPPPPELLQTLTAVGGIGOLTAAKESGG	412
Db	14	LFQYELTLHRTQHQDHNVVTG	GALELLQQFLRTPPPPELLQTLTAVGGIGOLTAAKESGG	73
Qy	413	RSRGSIVELLTAGGGSSCPVL	SRKQKGKVLGEGEEALEDDES	472
Db	74	RSRGSIVELLTAGGGSSCPVL	SRKQKGKVLGEGEEALEDDES	133
Qy	473	ISGELAASSGYSTPGSAGHDII	TEQPSRQHTLQADSVDLASC	532
Db	134	ISGELAASSGYSTPGSAGHDII	TEQPSRQHTLQADSVDLASC	193
Qy	533	SSQYSAVPSDPAMDNDGTQ	ASSPISDSSQTTTEGPD	592
Db	194	SSQYSAVPSDPAMDNDGTQ	ASSPISDSSQTTTEGPD	253
Qy	593	QIGQPQDEDEATGILPDEA	SEAFNSSMALQQAHL	652
Db	254	QIGQPQDEDEATGILPDEA	SEAFNSSMALQQAHL	313
Qy	653	TEPGDOENKPCRIGD	IGQSTDDDSAPLVHC	712
Db	314	TEPGDOENKPCRIGD	IGQSTDDDSAPLVHC	373
Qy	713	LALSCVGAVALHPESFFSK	LYKVPDLDTTEYPEQY	772
Db	374	LALSCVGAVALHPESFFSK	LYKVPDLDTTEYPEQY	433
Qy	773	TLICSIILSRFRHVGDM	NGTIRTTGTNTFSLAD	832
Db	434	TLICSIILSRFRHVGDM	NGTIRTTGTNTFSLAD	493
Qy	833	VMSLCSSSYSELGLQII	TDVTLRNSYVWLVRTE	892
Db	494	VMSLCSSSYSELGLQII	TDVTLRNSYVWLVRTE	553

QY 893 AHHTGLLKQERVLNNVVIHLGDEDPVRHVAASLIRLVPKLFYKCDQOQADPVVAV 952
Db 554 AHHTGLLKQERVLNNVVIHLGDEDPVRHVAASLIRLVPKLFYKCDQOQADPVVAV 613
QY 953 ARDOSSVYLKLLMHETQPPSHFSVSTITRIYRGYNLLPSITDVTMENNLSRVIAAVSHEL 1012
Db 614 ARDOSSVYLKLLMHETQPPSHFSVSTITRIYRGYNLLPSITDVTMENNLSRVIAAVSHEL 673
QY 1013 ITSTTRALTFGCEALCCLLSTAFPVCINSLGWHCGVPPPLSASDESRSKCTVGMATMILT 1072
Db 674 ITSTTRALTFGCEALCCLLSTAFPVCINSLGWHCGVPPPLSASDESRSKCTVGMATMILT 733
QY 1073 LSSANFPDLDSAHQDALIIDAGNLLAASAPKSLRSSWASSEEEANPAATKQEEVWPALGDRA 1132
Db 734 LSSANFPDLDSAHQDALIIDAGNLLAASAPKSLRSSWASSEEEANPAATKQEEVWPALGDRA 793
QY 1133 LVPWVEQLFSLHLLKVINICAHVLDVAVGPAIKAALPSLTNPPSLSPITRRKKEKEPGEQ 1192
Db 794 LVPWVEQLFSLHLLKVINICAHVLDVAVGPAIKAALPSLTNPPSLSPITRRKKEKEPGEQ 853
QY 1193 ASVPLSPKKGSEASASROSQDTSFPGVTTSSSLGSGFYHLPSTYKLHDV/LKATHANYKVT 1252
Db 854 ASVPLSPKKGSEASASROSQDTSFPGVTTSSSLGSGFYHLPSTYKLHDV/LKATHANYKVT 913
QY 1253 LDQNSTEKFGGFLRSALDVLSQLILELATLQDIGKCVBEIILGYLKSCFSREPMATVCVQ 1312
Db 914 LDQNSTEKFGGFLRSALDVLSQLILELATLQDIGKCVBEIILGYLKSCFSREPMATVCVQ 973
QY 1313 QLLKTLFGTNLASQFDGLSSNPSKSGRAQRGLGSSSVRPGLYHYCFMAYPTHFTQALADA 1372
Db 974 QLLKTLFGTNLASQFDGLSSNPSKSGRAQRGLGSSSVRPGLYHYCFMAYPTHFTQALADA 1033
QY 1373 SLRNMVQAEQENDTSG 1388
Db 1034 SLRNMVQAEQENDTSG 1049

Search completed: January 28, 2003, 14:39:40
Job time : 63.0972 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:10:57 ; Search time 32.0317 Seconds
(without alignments)
1417.333 Million cell updates/sec

Title: US-09-904-987-7
Perfect score: 7892
Sequence: 1 MATLEKLMKAFESLKSFQOO.....DGLMASGRKSPQPVRLQSP 1543

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*

1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*

2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*

3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*

4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*

5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep:*

6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7821.5	99.1	3144	2	US-08-457-273B-42
2	7821.5	99.1	3144	3	US-08-556-419-21
3	7821.5	99.1	3144	4	US-09-041-886-15
4	7805.5	98.9	3144	1	US-08-246-982A-6
5	7805.5	98.9	3144	1	US-08-453-265-6
6	7145.5	90.5	3119	1	US-08-246-982A-16
7	7145.5	90.5	3119	1	US-08-453-265-16
8	7057	89.4	3118	2	US-08-457-273B-8
9	2967.5	37.6	589	4	US-09-041-886-31
10	2784.5	35.3	552	4	US-09-041-886-30
11	2676.5	33.9	530	4	US-09-041-886-29
12	2589.5	32.8	513	4	US-09-041-886-28
13	157.5	2.0	1004	4	US-08-916-352-2
14	156	2.0	311	4	US-09-179-558-66
15	155	2.0	2441	1	US-08-194-468-2
16	155	2.0	2441	3	US-08-961-739-2
17	155	2.0	2441	4	US-09-514-247A-8
18	154	2.0	2442	4	US-09-514-247A-10
19	153.5	1.9	788	2	US-08-918-914-4
20	153.5	1.9	2842	1	US-07-741-940-7
21	153.5	1.9	2842	1	US-08-289-548A-7
22	153.5	1.9	2842	1	US-08-452-654-7
23	153.5	1.9	2842	4	US-08-449-731-7
24	153.5	1.9	2843	1	US-08-452-655B-2
25	153.5	1.9	2843	1	US-08-452-655B-7
26	153.5	1.9	2843	3	US-08-450-582-2
27	153.5	1.9	2843	3	US-08-450-582-7

28 153.5 1.9 2973 2 US-08-821-355A-7 Sequence 7, Appli
29 153.5 1.9 2973 2 US-09-003-687A-7 Sequence 7, Appli
30 153.5 1.9 2973 4 US-09-136-605-7 Sequence 7, Appli
31 153 1.9 370 4 US-09-142-551A-4 Sequence 4, Appli
32 153 1.9 396 4 US-09-142-551A-3 Sequence 3, Appli
33 153 1.9 2137 4 US-09-134-001C-4463 Sequence 4463, Ap
34 152.5 1.9 2843 1 US-07-741-940-2 Sequence 2, Appli
35 152.5 1.9 2843 1 US-08-289-548A-2 Sequence 2, Appli
36 152.5 1.9 2843 1 US-08-452-654-2 Sequence 2, Appli
37 152.5 1.9 2843 2 US-08-370-235A-2 Sequence 2, Appli
38 152.5 1.9 2843 4 US-08-449-731-2 Sequence 2, Appli
39 151.5 1.9 1719 4 US-09-605-785-378 Sequence 378, App
40 151.5 1.9 1719 4 US-09-439-313-378 Sequence 378, App
41 151.5 1.9 1719 4 US-09-352-616A-378 Sequence 378, App
42 149.5 1.9 521 5 PCT-US93-08386-10 Sequence 10, Appli
43 149.5 1.9 653 4 US-09-186-276B-2 Sequence 2, Appli
44 149.5 1.9 653 4 US-08-842-445-2 Sequence 2, Appli
45 149.5 1.9 653 4 US-09-186-188B-2 Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-08-457-273B-42
; Sequence 42, Application US/08457273B
; Patent No. 5849995
; GENERAL INFORMATION:
; APPLICANT: Hayden, Michael
; APPLICANT: Lin, Biaoyang
; APPLICANT: Nasir, Jamal
; TITLE OF INVENTION: Mouse Model for Huntington's Disease and
; TITLE OF INVENTION: Related DNA Sequences
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Virginia Bennett
; STREET: PO Box 37428
; CITY: Raleigh
; STATE: No. 5849995th Carolina
; COUNTRY: US
; ZIP: 27627
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/457,273B
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Bennett, Virginia C.
; REGISTRATION NUMBER: 37,092
; REFERENCE/DOCKET NUMBER: 3477-85A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-854-1400
; TELEFAX: 919-854-1401
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3144 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-457-273B-42

Query Match 99.1%; Score 7821.5; DB 2; Length 3144;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1533; Conservative 0; Mismatches 0; Indels 3; Gaps 1;
Qy 1 MATLEKLMKAFESLKSF---QQQQQQQQQQQQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPPPPPPPQA 57
Sequence 7, Appli
Db 1 MATLEKLMKAFESLKSFQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPPPPPPPQA 60

QY 238 FGNFANDNEIKVLLKAFIANLKSSPTIRRTAAGSAVSICQHSRRTOYFYSWLLNVLLGL 297
Db 241 FGNFANDNEIKVLLKAFIANLKSSPTIRRTAAGSAVSICQHSRRTOYFYSWLLNVLLGL 300
QY 298 LVPVEDEHSTLLGLVLLRLYLVLPLLOQOVKDTSLKSGFVTRKEMEYSPSAEQLVQVY 357
Db 301 LVPVEDEHSTLLGLVLLRLYLVLPLLOQOVKDTSLKSGFVTRKEMEYSPSAEQLVQVY 360
QY 358 ELTLHHTQHDHNVVTALELLQQLFRTPPPELLQPLTAVGGIGQLTAKEESGGRSRG 417
Db 361 ELTLHHTQHDHNVVTALELLQQLFRTPPPELLQPLTAVGGIGQLTAKEESGGRSRG 420
QY 418 SIVELIAGGSSCPVLSRKQKGVLLGEEAELEDDSESRSVSSSALTASVKDEISGEL 477
Db 421 SIVELIAGGSSCPVLSRKQKGVLLGEEAELEDDSESRSVSSSALTASVKDEISGEL 480
QY 478 AASSGVSTPGSAGHDITTPQRSOHTLQADSVDLASCDLTSSATDGEDDEILSHSSQVS 537
Db 481 AASSGVSTPGSAGHDITTPQRSOHTLQADSVDLASCDLTSSATDGEDDEILSHSSQVS 540
QY 538 AVPSDPAMDNDGTOASSPISDSOTTTEGPDPSAVTPSPDSSEIVLDGTDNQVGLQIQGP 597
Db 541 AVPSDPAMDNDGTOASSPISDSOTTTEGPDPSAVTPSPDSSEIVLDGTDNQVGLQIQGP 600
QY 598 QDEDEATGILPDEASEAFRNSSMALQQAHLKNNKSHCRQPSDSSVDKFLVLRDEATEPGD 657
Db 601 QDEDEATGILPDEASEAFRNSSMALQQAHLKNNKSHCRQPSDSSVDKFLVLRDEATEPGD 660
QY 658 QENKPCRIKGDIGOSTDDASPLVHCVRLLSASFLITGCKNVLPDPDRVRSVKALALSC 717
Db 661 QENKPCRIKGDIGOSTDDASPLVHCVRLLSASFLITGCKNVLPDPDRVRSVKALALSC 720
QY 718 VGAVALHPESFSLKYKVPDLTTEYPEQYVSDILNIDHGDPOVGRGATATLTCGLTICS 777
Db 721 VGAVALHPESFSLKYKVPDLTTEYPEQYVSDILNIDHGDPOVGRGATATLTCGLTICS 780
QY 778 ILSRFRHVGDMGTIRLTGNTFSLADCIPLLRKTLKDESSVTCKLACTAVRNCVMSLC 837
Db 781 ILSRFRHVGDMGTIRLTGNTFSLADCIPLLRKTLKDESSVTCKLACTAVRNCVMSLC 840
QY 838 SSSYSELGLQILIDVLTLNSSYWLVRTELLETLAEIDFLVSLFLEAKAENLHRCAGHYT 897
Db 841 SSSYSELGLQILIDVLTLNSSYWLVRTELLETLAEIDFLVSLFLEAKAENLHRCAGHYT 900
QY 898 GLLKLQERYLNNVVIHLLGDEDPVRRHVAASLIRLVPKLFYKCDQGOADPVAVARQDS 957
Db 901 GLLKLQERYLNNVVIHLLGDEDPVRRHVAASLIRLVPKLFYKCDQGOADPVAVARQDS 960
QY 958 SYLLKLMHETOPPSHFVSSTITRIYRGYNLLPSITDVTMNNLSRVIAAVSHELITSTT 1017
Db 961 SYLLKLMHETOPPSHFVSSTITRIYRGYNLLPSITDVTMNNLSRVIAAVSHELITSTT 1020
QY 1018 RALTFCCEALCLSTAPVCIWISLGHGCVPPPLSASDESRSKCTVGMATMILLSSAW 1077
Db 1021 RALTFCCEALCLSTAPVCIWISLGHGCVPPPLSASDESRSKCTVGMATMILLSSAW 1080
QY 1078 FPLDLSAHQDALILAGNLLAASAPKSLRSSWASEEENPAATKQEVNVPALGDRLVPMV 1137
Db 1081 FPLDLSAHQDALILAGNLLAASAPKSLRSSWASEEENPAATKQEVNVPALGDRLVPMV 1140
QY 1138 EQLFSLHLKLVINICAHVLDVAVGPAIKAALPSLTNPSPSLSPIRKKGKEPGEQASVPL 1197
Db 1141 EQLFSLHLKLVINICAHVLDVAVGPAIKAALPSLTNPSPSLSPIRKKGKEPGEQASVPL 1200
QY 1198 SPKKGSEASASRQSDTSGPVTTSKSSISGSFYHLPSYKLKLDVILKATHYKVTLDLQN 1257
Db 1201 SPKKGSEASASRQSDTSGPVTTSKSSISGSFYHLPSYKLKLDVILKATHYKVTLDLQN 1260
QY 1258 STEKFGGLRSALDVLQSLLELATLQDICKVCEETILGYLKSFCSPREPMMATVCVQQLLKT 1317
Db 1261 STEKFGGLRSALDVLQSLLELATLQDICKVCEETILGYLKSFCSPREPMMATVCVQQLLKT 1320
QY 1318 LFGTNLASQFDGLSSNPSKSGRAQLGSSSVRPGLYHYCFMAPYTHFTQALADASLRNM 1377

Db 1321 LFGTNLASQFDGLSSNPSKSGRAQLGSSSVRPGLYHYCFMAPYTHFTQALADASLRNM 1380
QY 1378 VQAEQENDTSGWFDVLQKVSTQTKNTLTSVTKNRADKNAIHHNRILRFLPEPLVIKALKQVTT 1437
Db 1381 VQAEQENDTSGWFDVLQKVSTQTKNTLTSVTKNRADKNAIHHNRILRFLPEPLVIKALKQVTT 1440
QY 1438 TTCVOLQKQVLDLLAQLVQLRVNYCLLSDSQVFTGFVLKQFEYIEVGQFRESEAIIPNIF 1497
Db 1441 TTCVOLQKQVLDLLAQLVQLRVNYCLLSDSQVFTGFVLKQFEYIEVGQFRESEAIIPNIF 1500
QY 1498 FFLVLLSYERYHSHKQIIGIPKIIQICDGMASGRKA 1533
Db 1501 FFLVLLSYERYHSHKQIIGIPKIIQICDGMASGRKA 1536

RESULT 3
US-09-041-886-15
; Sequence 15 Application US/09041886
; Patent No. 6235872
; GENERAL INFORMATION:
; APPLICANT: Bredesen, Dale E.
; APPLICANT: Rabizadeh, Sharroz
; TITLE OF INVENTION: Proapoptotic Peptides, Dependence
; TITLE OF INVENTION: Polypeptides and Methods of Use
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/041,886
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 2626
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3144 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-041-886-15

Query Match 99.1%; Score 7821.5; DB 4; Length 3144;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1533; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 1 MATLEKLMKAFESLKSP- --QQ 57
Db 1 MATLEKLMKAFESLKSP- --QQ 60
QY 58 QPLLPQPP 117
Db 61 QPLLPQPP 120
QY 118 PEQKLLGIAMELFLCSDDAESDVRVMADECLNKVTKALMDSNLPRLQLELYKEIKNG 177
Db 121 PEQKLLGIAMELFLCSDDAESDVRVMADECLNKVTKALMDSNLPRLQLELYKEIKNG 180

QY 178 APRSLAALWRAELAHLPVKRCRYLVNLLPCLTRTSKRPEESVQETLAAAVPKIMAS 237
DB 181 APRSLAALWRAELAHLPVKRCRYLVNLLPCLTRTSKRPEESVQETLAAAVPKIMAS 240
QY 238 FGFANDNEIKVLLKAFIANLKSSPTIRRTAAGSAVSCQHSRRRTQYFYSWLLNVLLGL 297
DB 241 FGFANDNEIKVLLKAFIANLKSSPTIRRTAAGSAVSCQHSRRRTQYFYSWLLNVLLGL 300
QY 298 LVPVEDEHSTLLTGLVLLRLYLVLVLLQOVKDTSLKSGFVTRKEMEVSPSAEOLVQVY 357
DB 301 LVPVEDEHSTLLTGLVLLRLYLVLVLLQOVKDTSLKSGFVTRKEMEVSPSAEOLVQVY 360
QY 358 ELTLHTHQHNDHNVVTALELLQQLFRTPELLOTLTAVGGTIGQLTAKEESGGRSRSG 417
DB 361 ELTLHTHQHNDHNVVTALELLQQLFRTPELLOTLTAVGGTIGQLTAKEESGGRSRSG 420
QY 418 SIVELIAGGSSCPVLSRKQKGVLLGEEALEDDSESVDVSSSALTASVKDEISGEL 477
DB 421 SIVELIAGGSSCPVLSRKQKGVLLGEEALEDDSESVDVSSSALTASVKDEISGEL 480
QY 478 AASSGVSTPGSAGHDITTEQPRSOHTLQADSVDLASCDLTSSATDGDDEEDILSHSSQVS 537
DB 481 AASSGVSTPGSAGHDITTEQPRSOHTLQADSVDLASCDLTSSATDGDDEEDILSHSSQVS 540
QY 538 AVPSDPAMDNDGQTASSPTSDSQTTEGPDSAVTPSDSSEIVLDGTDNQYGLQIGQP 597
DB 541 AVPSDPAMDNDGQTASSPTSDSQTTEGPDSAVTPSDSSEIVLDGTDNQYGLQIGQP 600
QY 598 QDEDEATGILPDEASEAFNNSMALQQAHLKKNMCHRCQPSDSSVDKFLVLRDEATEPGD 657
DB 601 QDEDEATGILPDEASEAFNNSMALQQAHLKKNMCHRCQPSDSSVDKFLVLRDEATEPGD 660
QY 658 QENKPRKIKGIDGSDSDSAPLVHCVRLLSASFLLTGKKNVLVDPDRVRSVKALALSC 717
DB 661 QENKPRKIKGIDGSDSDSAPLVHCVRLLSASFLLTGKKNVLVDPDRVRSVKALALSC 720
QY 718 VGAVALHPESFFSKLYKVPDLTTEPEQYVSDIILNYIDHGDPPQVRGATAILCGFLICS 777
DB 721 VGAVALHPESFFSKLYKVPDLTTEPEQYVSDIILNYIDHGDPPQVRGATAILCGFLICS 780
QY 778 ILSRPRHVDGNGMTIRLTGNTFSLADCIPLLRKTLKDESSVTKLACTAVRNCVMSLC 837
DB 781 ILSRPRHVDGNGMTIRLTGNTFSLADCIPLLRKTLKDESSVTKLACTAVRNCVMSLC 840
QY 838 SSSVSELGLQIIDLVLTLRNSVWLVRTELETLAEIDFRLVSFLAKAENLHRGAHHT 897
DB 841 SSSVSELGLQIIDLVLTLRNSVWLVRTELETLAEIDFRLVSFLAKAENLHRGAHHT 900
QY 898 GLKLQERVLNNVYIHLGSDDEPRVRHVAASLIRLVKLFYKCDQGDQADPVVAVARDQS 957
DB 901 GLKLQERVLNNVYIHLGSDDEPRVRHVAASLIRLVKLFYKCDQGDQADPVVAVARDQS 960
QY 958 SVYLLKLMHETQPPSHFSVSTIIRIRGVNLLPSITDVTMNNLSRVIAAVSHELITSTT 1017
DB 961 SVYLLKLMHETQPPSHFSVSTIIRIRGVNLLPSITDVTMNNLSRVIAAVSHELITSTT 1020
QY 1018 RALTEGCEALCLLSTAFFVCIWLSGHCVPPLSASDESRSKCTVGMATMILTLLSSAW 1077
DB 1021 RALTEGCEALCLLSTAFFVCIWLSGHCVPPLSASDESRSKCTVGMATMILTLLSSAW 1080
QY 1078 FPLDLSAHQDALILAGNLLAASAPKSLRSSWASEEENPAATKOEVEWPAIGDRALVPMV 1137
DB 1081 FPLDLSAHQDALILAGNLLAASAPKSLRSSWASEEENPAATKOEVEWPAIGDRALVPMV 1140
QY 1138 EQLFESHLKLVINTICAHVLDVAVGPAKALPSLTNPSPSPIRRGKEKEPEQASVPL 1197
DB 1141 EQLFESHLKLVINTICAHVLDVAVGPAKALPSLTNPSPSPIRRGKEKEPEQASVPL 1200
QY 1198 SPKKGSEASARSQSDTSGPVTTSSKSSLSGFSYHLPSYKLHDLVKATHANYKVTLDLQN 1257
DB 1201 SPKKGSEASARSQSDTSGPVTTSSKSSLSGFSYHLPSYKLHDLVKATHANYKVTLDLQN 1260
QY 1258 STEKFGGFLRSALDVLSQLLELATLQDIGKCVBEILGYLKSCFSREPMMATVCVQQLLKT 1317

DB 1261 STEKFGGFLRSALDVLSQLLELATLQDIGKCVBEILGYLKSCFSREPMMATVCVQQLLKT 1320
QY 1318 LFGFNLASQFGLSSNPSKSGRAQLGSSSVRPLGYHYCFMAYTHFTQALADASLRNM 1377
DB 1321 LFGFNLASQFGLSSNPSKSGRAQLGSSSVRPLGYHYCFMAYTHFTQALADASLRNM 1380
QY 1378 VQAEQENDTSGWFQVLOKVSTQLKTNLTSVTKNRADKNAINHNHRLFEPLVIRKALKOYTT 1437
DB 1381 VQAEQENDTSGWFQVLOKVSTQLKTNLTSVTKNRADKNAINHNHRLFEPLVIRKALKOYTT 1440
QY 1438 TTCVOLQKQVLDLALQVLQVRVNYCLLSDSQVFTGFVLKQFEYTEVQFRESEAIIPNIF 1497
DB 1441 TTCVOLQKQVLDLALQVLQVRVNYCLLSDSQVFTGFVLKQFEYTEVQFRESEAIIPNIF 1500
QY 1498 FFLVLLSYERYHSKOIIGIPKIIQICOGIMASGRKA 1533
DB 1501 FFLVLLSYERYHSKOIIGIPKIIQICOGIMASGRKA 1536

RESULT 4

US-08-246-982A-6
; Sequence 6, Application US/08246982A
; Patent No. 5686288
; GENERAL INFORMATION:
; APPLICANT: MacDonald, Marcy E.
; APPLICANT: Ambrose, Christine M.
; APPLICANT: Duyao, Mabel P.
; APPLICANT: Gusella, James F.
; TITLE OF INVENTION: Huntingtin DNA, Protein And Uses Thereof
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Steirne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: May 20, 1994
; APPLICATION NUMBER: US/08/246.982A
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldstein, Jorge, A.
; REGISTRATION NUMBER: 29,021
; REFERENCE/DOCKET NUMBER: 0609.3880002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3144 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-246-982A-6

Query Match 98.9%; Score 7805.5; DB 1; Length 3144;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1530; Conservative 2; Mismatches 1; Indels 3; Gaps 1;

QY 1 MATLEKLMKAFESLSKF---QQQQQQQQQQQQQQQQQQQQPPPPPPPPPPQLPQPQA 57
DB 1 MATLEKLMKAFESLSKFQQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPPPPPQA 60
QY 58 QPLLPOQPP 117
DB 61 QPLLPOQPP 120

QY 118 PEFQKLLGIAMELFLLCSDDAESDVRMADECLNKVICALMDSNLRLOLELYKEIKKNG 177
DB 121 PEFQKLLGIAMELFLLCSDDAESDVRMADECLNKVICALMDSNLRLOLELYKEIKKNG 180
QY 178 APRSLAALRFRFAELAHVLPQCRPYLVNLLPCLRTSKRPESVOETLAAAVPKIMAS 237
DB 181 APRSLAALRFRFAELAHVLPQCRPYLVNLLPCLRTSKRPESVOETLAAAVPKIMAS 240
QY 238 FGFANDNEIKVLLKAFIANLKSSPTIRRTAAGSAVSIQHSRRTOYFYSWLLNVLGL 297
DB 241 FGFANDNEIKVLLKAFIANLKSSPTIRRTAAGSAVSIQHSRRTOYFYSWLLNVLGL 300
QY 298 LVPVEDEHSTLLILGVLTLRLYLPLVLLQQVKDTSLKSGFVTRKEMEVSPSAEQLVQVY 357
DB 301 LVPVEDEHSTLLILGVLTLRLYLPLVLLQQVKDTSLKSGFVTRKEMEVSPSAEQLVQVY 360
QY 358 ELTLHHTQHODHNVVVGALQLLQQLFRTPPPELLQTLTAVGGTGQJTAKEESGGRRSG 417
DB 361 ELTLHHTQHODHNVVVGALQLLQQLFRTPPPELLQTLTAVGGTGQJTAKEESGGRRSG 420
QY 418 SIVELIAGGSGSPVLSRKQKGVLLGEEALEDDSESDVSSSALTASVKDEISGEL 477
DB 421 SIVELIAGGSGSPVLSRKQKGVLLGEEALEDDSESDVSSSALTASVKDEISGEL 480
QY 478 AASSGVSTPGSAGHDIITEPRSQHTLQADSVDLASCDLTSSATDDEEDILSHSSSQVS 537
DB 481 AASSGVSTPGSAGHDIITEPRSQHTLQADSVDLASCDLTSSATDDEEDILSHSSSQVS 540
QY 538 AVPSDPAMDNDGTQASSPISDSQTTTEGPDSAVTPSDSSSEIVLDGTDNQYLGLQIGQP 597
DB 541 AVPSDPAMDNDGTQASSPISDSQTTTEGPDSAVTPSDSSSEIVLDGTDNQYLGLQIGQP 600
QY 598 QDEDEATGLPDEAFAPNSMAQQOHLKKNMHCRCQPSDSSVDKFLVRDEATEPGD 657
DB 601 QDEDEATGLPDEAFAPNSMAQQOHLKKNMHCRCQPSDSSVDKFLVRDEATEPGD 660
QY 658 QENKPCRIKIDIGQSTDDSDAPLVHCVRLTSLASFLTGTGKNVLPDRDVRVSVKALALSC 717
DB 661 QENKPCRIKIDIGQSTDDSDAPLVHCVRLTSLASFLTGTGKNVLPDRDVRVSVKALALSC 720
QY 718 VGAVALHPESFSSKLYKVPDTEPEQYVSDILNYIDHGDPQVRGATAILCGTFLICS 777
DB 721 VGAVALHPESFSSKLYKVPDTEPEQYVSDILNYIDHGDPQVRGATAILCGTFLICS 780
QY 778 ILRSRPHVGDWGTIRTLTGNTFSLADCIPLLRKTLKDESSVTCKLACTAVRNCVMSLC 837
DB 781 ILRSRPHVGDWGTIRTLTGNTFSLADCIPLLRKTLKDESSVTCKLACTAVRNCVMSLC 840
QY 838 SSSYSELGLQIIDLVLTLRNSYVLRTELLETLAETDFRLVSFLAKAENLHRGAHHT 897
DB 841 SSSYSELGLQIIDLVLTLRNSYVLRTELLETLAETDFRLVSFLAKAENLHRGAHHT 900
QY 898 GLLKQERLVNANNVHLLGDEDEPRVRHVAASLIRLVKFLYKCDQGDQADPVVAVARDQS 957
DB 901 GLLKQERLVNANNVHLLGDEDEPRVRHVAASLIRLVKFLYKCDQGDQADPVVAVARDQS 960
QY 958 SVYKLLMHETQPPSHFSVTIRYRGYNLPSITDVTMNNLSRVIAAVSHELITST 1017
DB 961 SVYKLLMHETQPPSHFSVTIRYRGYNLPSITDVTMNNLSRVIAAVSHELITST 1020
QY 1018 RALTFCCEALCLLSTAFFVCIVSLGWHCVPPLPSADESRKSTCTVGMATMILTLLSSAW 1077
DB 1021 RALTFCCEALCLLSTAFFVCIVSLGWHCVPPLPSADESRKSTCTVGMATMILTLLSSAW 1080
QY 1078 FPLDLAHDQDALILAGNLLAASAPKSLRSSWASEEENPAATKQEEVWPALGDRLVPMV 1137
DB 1081 FPLDLAHDQDALILAGNLLAASAPKSLRSSWASEEENPAATKQEEVWPALGDRLVPMV 1140
QY 1138 EQLFSLHLLKVINICAHVLDVAPGPAKALPSLTNPSPSPTRRRKGKEPEGEQASVPL 1197
DB 1141 EQLFSLHLLKVINICAHVLDVAPGPAKALPSLTNPSPSPTRRRKGKEPEGEQASVPL 1200

QY 1198 SPKKGSEASAASRQSDTSGPVTTSSKSSSLGSPYHLPVSLKHLHDVLKATHANYKVTLDLQN 1257
DB 1201 SPKKGSEASAASRQSDTSGPVTTSSKSSSLGSPYHLPVSLKHLHDVLKATHANYKVTLDLQN 1260
QY 1258 STEKEGGFURSALDVLSQLLELATLQDYGKVEEILGVLKSCFSREPMMATVCVQOLKLT 1317
DB 1261 STEKEGGFURSALDVLSQLLELATLQDYGKVEEILGVLKSCFSREPMMATVCVQOLKLT 1320
QY 1318 LFGTNLASQFDCGLSNPSKQGRAGRLGSSSVRPGLYHYCFMAYTHFTQALADASLRNM 1377
DB 1321 LFGTNLASQFDCGLSNPSKQGRAGRLGSSSVRPGLYHYCFMAYTHFTQALADASLRNM 1380
QY 1378 VQAEQENDTSGWFDVLQKVSTQKLTNLTSVTKNRADKNAIHNIHLRFLPVLTKALKQVTT 1437
DB 1381 VQAEQENDTSGWFDVLQKVSTQKLTNLTSVTKNRADKNAIHNIHLRFLPVLTKALKQVTT 1440
QY 1438 TTCVOLQKQVLDLLAQLVOLRVNYCLLSDQVFIQGVFLKQFEYIEVGQFRESEAIIPNIF 1497
DB 1441 TTCVOLQKQVLDLLAQLVOLRVNYCLLSDQVFIQGVFLKQFEYIEVGQFRESEAIIPNIF 1500
QY 1498 FFLVLLSVERYHSKQIIGIPKIIQICDGMASGRKA 1533
DB 1501 FFLVLLSVERYHSKQIIGIPKIIQICDGMASGRKA 1536

RESULT 5

US-08-453-265-6
; Sequence 6, Application US/08453265
; Patent No. 5693757
; GENERAL INFORMATION:
; APPLICANT: MacDonald, Marcy E.
; APPLICANT: Ambrose, Christine M.
; APPLICANT: Duyao, Mabel P.
; APPLICANT: Gusella, James F.
; TITLE OF INVENTION: Huntingtin DNA, Protein And Uses Thereof
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/453,265
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3880003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3144 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-453-265-6

Query Match 98.9%; Score 7805.5; DB 1; Length 3144;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1530; Conservative 2; Mismatches 1; Indels 3; Gaps 1;
QY 1 MATLEKLMKAPSLKSF---QQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPPPQA 57
|||||

[illegible]

QY 298 LVPVEDEHSTLLILGVLRLYLVLPLQQQVKDTSLKSGFVTRKEMEVSPPSAEQLVQVY 357
 Db 301 LVPVEDEHSTLLILGVLRLYLVLPLQQQVKDTSLKSGFVTRKEMEVSPPSAEQLVQVY 360
 QY 358 ELTLHTHQDHNVVTGALQLLQQLFRTPPPPELLOTLTAVGGIGQLTAKEESGGRSRSG 417
 Db 361 ELTLHTHQDHNVVTGALQLLQQLFRTPPPPELLOTLTAVGGIGQLTAKEESGGRSRSG 420
 QY 418 SIVELTAGGSSCPVLSRKQKGVLLGEEALEDDSESRSVSSSALTASVKDEISGEL 477
 Db 421 SIVELTAGGSSCPVLSRKQKGVLLGEEALEDDSESRSVSSSALTASVKDEISGEL 480
 QY 478 AASSGVSTPGSAGHDIIITEQPSQHTLQADSVD 510
 Db 481 AASSGVSTPGSAGHDIIITEQPSQHTLQADSVD 513

RESULT 13

US-08-916-352-2
 ; Sequence 2, Application US/08916352
 ; Patent No. 6166191
 ; GENERAL INFORMATION:
 ; APPLICANT: CHIRON CORPORATION
 ; TITLE OF INVENTION: HUMAN POLYHOMEOTIC 1 (hph1) ACTS AS A
 ; TITLE OF INVENTION: TUMOR SUPPRESSOR
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: CHIRON CORPORATION
 ; STREET: 4560 HORTON STREET
 ; CITY: EMERYVILLE
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94608
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/916.352
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: POTTER, JANE
 ; REGISTRATION NUMBER: 33,332
 ; REFERENCE/DOCKET NUMBER: 1355.
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 510-923-2707
 ; TELEFAX: 510-655-3542
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1004 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-916-352-2

Query Match 2.0%; Score 157.5; DB 4; Length 1004;
 Best Local Similarity 39.1%; Pred. No. 0.002;
 Matches 45; Conservative 8; Mismatches 33; Indels 29; Gaps 5;

QY 10 AFESLSKSFQQQQQ-----QQQQQQQQQQQQQQQ-----PPPPPPPPPPQL 50
 Db 400 AIHQQQQFQHQSGQLLHTATHLQAAQQQQQQQQQQQQQQAATLTAPQPPQVPTQV 459
 QY 51 P-QPPQAQPLLPQ-----QPPPPPPPPPPGPAVAEPLHRPK-KELSATK 95
 Db 460 PPSQSQQAQTLVQPMQLSSPLSLPPDAAPKPPPIQSKPPVAPKPKPOLCAAK 514

RESULT 14

US-09-179-558-66

; Sequence 66, Application US/09179558
 ; Patent No. 6180612
 ; GENERAL INFORMATION:
 ; APPLICANT: Hockensmith, Joel W.
 ; APPLICANT: Muthuswami, Rohini
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
 ; TITLE OF INVENTION: TARGETING DNA METABOLIC PROCESSES USING
 ; TITLE OF INVENTION: AMINOGLYCOSIDE DERIVATIVES
 ; NUMBER OF SEQUENCES: 66
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: PENNIE & EDMONDS LLP
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: NY
 ; COUNTRY: USA
 ; ZIP: 10036-2711
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/179,558
 ; FILING DATE: 27-OCT-1998
 ; CLASSIFICATION: 514
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: U.S. 09/060,470
 ; FILING DATE: 15-APR-1998
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: U.S. 60/063,898
 ; FILING DATE: 31-OCT-1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Coruzzi, Laura A
 ; REGISTRATION NUMBER: 30,742
 ; REFERENCE/DOCKET NUMBER: 9426-005-999
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212)7909090
 ; TELEFAX: (212)8699741
 ; TELEX: 66141 PENNIE
 ; INFORMATION FOR SEQ ID NO: 66:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 311 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-09-179-558-66

Query Match 2.0%; Score 156; DB 4; Length 311;
 Best Local Similarity 23.8%; Pred. No. 0.0004;
 Matches 91; Conservative 34; Mismatches 116; Indels 142; Gaps 20;

QY 38 PPPPPPPPPPPQLPQQPPQA-----QPLLPOQ-PPPPPPPPPPGPAVAEPLHRPK 88
 Db 4 PPP 62
 QY 89 KELSATKKDRVNHCLTICENIVAQSVRNSDEFQKLLGIAMELFLLCSDDAESDVRVADE 148
 Db 63 -ELFA-----SGALTPEPFKATFSQ--FPKSDQGGSPS---PDQ 98
 QY 149 CLNKVIKALMDSNLPRLQLELYKEIKKNGAPRSRLAARFAELAHVLRPKCRPYLVNL 208
 Db 99 C-----QDQTPP-----GSPRRRX-----PPTK----- 118
 QY 209 LPLTRTSKRPE-ESVQETLAAVPM-----ASFGNFANDNEIKVLLKAFIA 256
 Db 119 ---LSSSKKKFSFSFNRTADEAKIPSTIEYILDLESGREKFLVFAHH---KVVLDAITK 172
 QY 257 NLKSSSPITIR---TAAGSAVSICQHSRRRTQFYFYSMLLVNVL---LGLLVPEDEHSTLL 309
 Db 173 ELEKRVQHIRIDGSTSSADREDLCQQQLSPGPAVALSITANMGLIFSSAD----LV 228
 QY 310 ILGLV-----LTLR-----YLVLQOQVQKDTSLKG-- 335

